integrated systems inc.

2500 Mission College Boulevard
Santa Clara, California 95054-1215
Telex 559631 • FAX: (408) 980-0400

Annual Technical Report
1 June 1987 through 30 Sept. 1988

Adaptive Control Techniques for Large Space Structures

(Contract No. F49620-85-C-0094)

Prepared by:

Dr. Robert L. Kosut

Integrated Systems Inc.
2500 Mission College Boulevard
Santa Clara, California 95054-1215

Prepared for:

AFOSR, Directorate of Aerospace Sciences
Bolling Air Force Base
Washington, DC 20332

Attention:

Dr. Anthony Amos
Building 410

ISI Report No. 150
6 Jan. 1989

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

ADA204530

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION
Unclassified

1b. RESTRICTIVE MARKINGS

2a. SECURITY CLASSIFICATION AUTHORITY

2b. DECLASSIFICATION/DOWNGRADING SCHEDULE

3. DISTRIBUTION/AVAILABILITY OF REPORT
Approved for Public Release;
Distribution is unlimited

4. PERFORMING ORGANIZATION REPORT NUMBER(S)
ISI No.150

5. MONITORING ORGANIZATION REPORT NUMBER(S)
AFOSR TR 0071

6a. NAME OF PERFORMING ORGANIZATION
INTEGRATED SYSTEMS, INC.

6b. OFFICE SYMBOL *(If applicable)*

7a. NAME OF MONITORING ORGANIZATION
AFOSR/NA

6c. ADDRESS *(City, State and ZIP Code)*
2500 Mission College Blvd.
Santa Clara, CA 95054

7b. ADDRESS *(City, State and ZIP Code)*
Building 410, Bolling AFB, DC 20332-6448

8a. NAME OF FUNDING/SPONSORING ORGANIZATION
DCASMA

8b. OFFICE SYMBOL *(If applicable)*
NA

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
F49620-85-C-0094

8c. ADDRESS *(City, State and ZIP Code)*
[redacted] same as 7b.

10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
61102F	2302	B1	

11. TITLE *(Include Security Classification)*
Adaptive Control Techniques for Large Space Structures (U)

12. PERSONAL AUTHOR(S)
Robert L. Kosut, Michael G. Lyons

13a. TYPE OF REPORT
Final Technical

13b. TIME COVERED
FROM 6/1/87 TO 9/30/88

14. DATE OF REPORT *(Yr., Mo., Day)*
1/6/89

15. PAGE COUNT

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB. GR.

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*
Adaptive Control, On-line Adaptation,
Nonlinear Flexible Systems

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

This report summarizes the research performed during the period 1 June 1987 to 30 Sept 1988. on adaptive control techniques for Large Space Structures (LSS). The research effort concentrated on two areas: (1) on-line robust design from identified models - what is referred to here as uncertainty estimation, and (2) slow adaptive control of nonlinear flexible systems.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT
UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☒ DTIC USERS ☒

21. ABSTRACT SECURITY CLASSIFICATION
Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL
Dr Anthony K. Amos

22b. TELEPHONE NUMBER *(Include Area Code)*
(202) 767-4937

22c. OFFICE SYMBOL
AFOSR/NA

DD FORM 1473, 83 APR

EDITION OF 1 JAN 73 IS OBSOLETE.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE

Contents

Chapter 1

Introduction

1.1 Background and Motivation

The Large Space Structure (LSS) research program was originally formulated in late 1982 in response to the increasing concern that performance robustness of Air Force LSS type system would be inadequate to meet mission objectives. In particular, uncertainties in both system dynamics and disturbance spectra characterizations (both time varying and stochastic uncertainty) significantly limit the performance attainable with fixed gain, fixed architecture controls. Therefore, the use of an adaptive system, where disturbances and/or plant models are identified prior to or during control, gives systems designers more options for minimizing the risk in achieving performance objectives.

The aim of adaptive control is to implement in real-time and on-line as many as possible of the design functions now performed off-line by the control engineer; to give the controller "intelligence". To realize this aim, both a theory of stability and performance of such inherently nonlinear controls is essential as well as a technology capable of achieving the implementation.

The issues of performance sensitivity, robustness, and achievement of very high performance in an LSS system can be effectively addressed using adaptive algorithms. The need to identify modal frequencies, for example, in high-performance disturbance rejection systems has been shown in ACOSS (1981) and VCOSS (1982). The deployment of high-performance optical or RF systems may require on-line identification of critical modal parameters before full control authority can be exercised. Parameter sensitivity, manifested by performance degradation or loss of stability (poor robustness) may be effectively reduced by adaptive feedback mechanizations. Reducing the effects of on-board disturbance rejection) is particularly important for planned Air Force missions. For these cases, adaptive control mechanizations are needed to produce the three-to-five orders-of-magnitude reductions in line-of-sight jitter required by the mission.

Research is essential to identify the performance limitations of adaptive strategies for LSS control both from theoretical and hardware mechanization viewpoints. The long range goal of this research program is to establish guidelines for selecting the appropriate strategy, to evaluate performance improvements over fixed-gain mechanizations, and to examine the architecture necessary to produce a practical hardware realization. The initial thrust, however, is to continue to build a strong theoretical foundation without losing sight of the practical implementation issues.

1.2 Research Objectives

The aims of this research study are to extend and develop adaptive control theory and its application to LSS in several directions. These include:

1. **Theoretical Development:** The initial emphasis has been on slow adaptation, since this covers may LSS situations. Later on we will examine fast adaptation. The theory developed here will provide for:
   (a) estimates of robustness, i.e., stability margins vs. performance bounds;
   (b) estimates of regions of attraction and rates of parameter convergence to these regions;
   (c) extension of the present linear finite dimensional adaptive theory to include nonlinear and infinite dimensional plants and controller structures; and
   (d) extensions to decentralized systems.
2. **Parameter Adaptive Algorithms** Assess the behavior of different algorithms, including: gradient, recursive least squares, normalized least mean squares, and nonlinear observer (e.g, Extended Kalman Filter).
3. **Parametric Models:** Assess the impact of model choices. In particular we will examine the effect of explicit and implicit model choices. An explicit model, for example, is a transfer function whose coefficients are all unknown. In an implicit model transfer function, the coefficients would be functions of some other parameters. Implicit models usually arise from physical or experimental data, whereas explicit models are selected for analytical convenience.
4. **Adaptive Nonlinear Control:** Although our early effort is to study adaptive linear control, there are may LSS situations where the control is nonlinear, e.g., large angle maneuvers, slewing.

1.3 Current Status

At the present time we stand at the beginning stages of the theoretical development in adaptive control. The result of recent efforts are contained in the selected papers in the Appendix and the references therein. A summary of earlier efforts is contained in the recently published textbook *Stability of Adaptive Systems: Passivity and Averaging Analysis*, MIT Press, 1986. This publication is an outgrowth of research supported under this contract and involved a considerable amount of collaborative effort among several researchers in the field of adaptive control. The text discusses adaptive systems from the viewpoint of stability theory. The emphasis is on methodology and basic concepts, rather than on details of adaptive algorithm. The analysis reveals common properties including causes and mechanisms for instability and the means to counteract them. Conditions for stability are presented under slow adaptation, where the method of averaging is utilized. In this latter case the stability result is *local*, i.e., the initial parametrization and input spectrum is constrained. Based on this analysis, a conceptual framework is now available to pursue the issues of slow adaptive control of LSS.

To remove the restrictiveness of slow adaptation requires an understanding of the transient behavior of adaptive systems. A preliminary investigation is reported in Kosut et al (1986). The transient behavior of not-slow or even rapid adaptation is a significant problem in the adaptive control of LSS, e.g., rapid retargeting.

Another approach to adaptive control is to calibrate (or tune) the controller based on a current estimate of the LSS model. This involves not just knowing *one* model, but rather, a model *set*. This problem, which we refer to as *adaptive calibration*, is essentially that of developing a technique of on-line robust control design from an identified model. Although we have worked on this problem for some time it is only recently that we have established a theoretical basis for estimating model error from system identification, Kosut(1987,1988), and this Appendix. This research has raised many new questions which need to be considered, e.g., what is the appropriate robust controller parametrization; how does it relate to model parametrization; how to iterate on the data if the estimate of model error is too large; what are the heuristics for experiment design.

1.4 Organization of Report

Throughout the remainder of the report we will discuss our activities this past year in uncertainty estimation and adaptive nonlinear control.

The Appendix contains a preprint of an extensive journal article on uncertainty estimation, in particular, on the subject of parameter set estimation. There still remains much to be done in this area, particularly in regard to control design using models with uncertain parameters. Another major effort in this area is in the use of fractional representations to aid in system identification in closed-loop. As the work is still in progress, we postpone discussion of this subject for the next report, although a conference paper will be presented in 1989 (see below). There are now many researchers interested in this approach, and two invited sessions have been organized for the 1989 ACC.

In Chapter 2 we discuss a new research direction, namely, the adaptive control of nonlinear flexible systems. Here we encounter unique nonlinear phenomena such as limit cycles, bifurcations, and chaos. An example is presented which demonstrates adaptation of a chaotic system. This area is still very new, thus, much of our discussion is in the future tense, that is, what we hope can be done in future research.

1.5 Selected Publications to Date

1.5.1 Journals and Conferences

R.L. Kosut, "Adaptive Control via Parameter Set Estimation", to appear, *Int. J. of Adaptive Control and Signal Processing*, 1989.

R.L. Kosut, "Adaptive Control of a Chaotic System", to appear, Proc. 1988 ACC, Pittsburg, PA, June 1989.

F. Hanson, R.L. Kosut, and G.F. Franklin, "Closed-Loop System Identification via the Fractional Representation: Experiment Design and Unmodeled Dynamics" , to appear, Proc. 1988 ACC, Pittsburg, PA, June 1989.

R.L. Kosut, "On The Use of The Method of Averaging for the Stability Analysis of Adaptive Linear Control Systems", *Proc. IEEE CDC*, Los Angeles, CA, Dec. 1987.

R.L. Kosut, "Conditions for Convergence and Divergence of Parameter Adaptive Linear Systems", *Proc. ISCAS 1987*, Philda., PA, May, 1987.

R.L. Kosut, "Adaptive Calibration: An approach to Uncertainty Modeling and On-Line Robust Control Design", *Proc. 25th IEEE CDC*, Athens, Greece, Dec. 1986.

R.L. Kosut, "Towards an On-Line Procedure for Automated Robust Control Design: The Adaptive Calibration Problem", presented at *1986 ACC*, June 1986.

R.L. Kosut,[1] I.M.Y. Mareels, B.D.O. Anderson, R.R. Bitmead, and C.R. Johnson, Jr., "Transient Analysis of Adaptive Control", submitted to *IFAC 10th World Congress*, Munich, Germany, July 1987.

R.L. Kosut,[1] and R.R. Bitmead, "Fixed-Point Theorems for Stability Analysis of Adaptive Systems", *Proc. IFAC Workshop on Adaptive Systems*, Lund, Sweden, July 1986.

R.L. Kosut,[1] and R.R. Bitmead, "Linearization of Adaptive Systems: A Fixed-Point Analysis", submitted, *IEEE Trans. on Circuits and Systems; Special Issue on Adaptive Systems*, Sept. 1987.

I.M.Y. Mareels, R.R. Bitmead, M. Gevers, C.R. Johnson, Jr., R.L. Kosut,[1] and M.A. Poubelle, "How Exciting Can a Signal Really Be?", to appear, *Systems and Control Letters*.

R.L. Kosut,[1] B.D.O. Anderson, and I.M.Y. Mareels, "Stability Theory for Adaptive Systems: Methods of Averaging and Persistency of Excitation", *IEEE Trans. on Aut. Contr.*, to appear, Jan. 1987.

B.D.O. Anderson, R.R. Bitmead, C.R. Johnson, Jr., and R.L. Kosut, "Stability Theorems for the Relaxation of the SPR Condition in Hyperstable Adaptive Systems", *IEEE Trans. on Aut. Contr.*, submitted.

R.L. Kosut and C.R. Johnson, Jr., "An Input-Output View of Robustness in Adaptive Control", *Automatica: Special Issue on Adaptive Control*, **20**(5):569-581, Sept. 1984.

R.L. Kosut,[2] and B. Friedlander, "Robust Adaptive Control: Conditions for Global Stability", *IEEE Trans. on Aut. Contr.*, **AC-30**(7):610-624, July 1985.

1.5.2 Books

B.D.O. Anderson, R.R. Bitmead, C.R. Johnson, Jr., P.V. Kokotovic, R.L. Kosut, I.M.Y. Mareels, L. Praly, and B.D. Riedle, *Stability of Adaptive Systems: Passivity and Averaging Analysis*, MIT Press, 1986.

R.L. Kosut, "Methods of Averaging for Adaptive Systems", *Adaptive Systems: Theory and Applications*, Editor: K.S. Narendra, Plenum Press, 1986.

R.L. Kosut and M.G. Lyons, "Issues in Control Design for Large Space Structures", *Adaptive Systems: Theory and Applications*, Editor: K.S. Narendra, Plenum Press, 1986.

[1] Research performed while R.L. Kosut was a Visiting Fellow at the Australian National University.

[2] Started under contract F4920-81-C-0051.

R.L. Kosut, "Adaptive Control of Large Space Structures: Uncertainty Estimation and Robust Control Calibration", *Large Space Structures: Dynamics and Control*, Editors: S.N. Atluri and A.K. Amos, Springer-Verlag, 1987.

E. Crawley, R.L. Kosut, and S. Hall, *Controlled Structures*, in preparation.

1.5.3 Other Related Publications

A.M Pascoal, R.L. Kosut, D. Meldrum, M. Workman, and G.F. Franklin, "Adaptive Time-Optimal Control of Flexible Structures", to appear, Proc. 1988 ACC, Pittsburg, PA, June 1989.

M.L. Workman, R.L. Kosut,[3] and Franklin, "Adaptive Proximate Time-Optimal Servomechanisms: Continuous-Time Case", *Proc. ACC*, pp.589-594, June 1987, Minneapolis, MN.

M.L. Workman, R.L. Kosut[3], and Franklin, "Adaptive Proximate Time-Optimal Servomechanisms: Discrete-Time Case", *Proc. CDC*, Dec. 1987, Los Angeles, CA.

R.L. Kosut,[4] A. Pascoal, S. Morrison, and M.L. Workman, "Time-Optimal Control of Large Space Structures", *Proc. SPIE*, Jan. 1988, Los Angeles, CA.

S. Philips, R.L. Kosut[3], and G.F. Franklin, "An Averaging Analysis of Discrete-Time Indirect Adaptive Control", to appear, *Proc. ACC*, June 1988, Atlanta, GA.

[3]Research supported partly by NSF Industry/ University Cooperative Research Program under Grant ECS-8605646.

[4]Resarch supported by SDIO/IST and managed by AFOSR.

Chapter 2

Overview of Research Activities

A new objective for this years work is the development of mathematical and computational tools for the analysis and synthesis of adaptive control of nonlinear flexible systems. With few exceptions, practically all of recent research in adaptive control is for the case of adaptive *linear* systems, *i.e.*, when the adaptive parameters are held fixed, the resulting system is linear and time-invariant. The research issues we discuss here arise not only from the nonlinearity introduced by the adaptation, but the nonlinearities present in the mechanical system itself. Both of these types of nonlinearities can produce bifurcations and chaotic effects. The principal aims of the this research include understanding the causes of these phenomena, how to prevent them from occurring, how to minimize any deleterious effects, and how to recognize them from measurements.

Nonlinearities in flexible mechanical systems arise from a number of sources including:

- actuator saturation
- friction in bearings or gear trains
- backlash due to gear spacing or slippage
- kinematic transformations
- aerodynamic forces

Each of these types of nonlinearities can influence the design of the controller. For example, efficient use of actuator authority in the presence of saturation motivates time-optimal control. Friction can be offset by either direct cancelation or dither. The effects of backlash can be handled with a controller that includes a pre-load. Nonlinearities arising from kinematic transformatioms or aerodynamic forces can be accomodated by using a controller that either incorporates gain scheduling as a function of equilibrium points, or else provides a "feedback linearization".

In a general way the control structure depends on the type of plant nonlinearity and some associated parameters. In many cases the type of nonlinearity as well as these parameter values are not very well known and prior knowledge is too coarse to guarantee acceptable closed-loop performance. Under these conditions, it is necessary to adapt the controller to existing conditions.

An adaptive control system can adjust to uncertainty by tuning controller parameters in such a way as to minimize model parameter inaccuracies as well as nullifying the effect of inaccurately (or even unknown) nonlinearities. There are many ways to design or configure an adaptive control system. A "generic" configuration for an adaptive control system is depicted in Figure 2.1. The control system consists of two processes which make it adaptive, namely: (1) a model parameter estimator (PE), and (2) a control design rule (D).

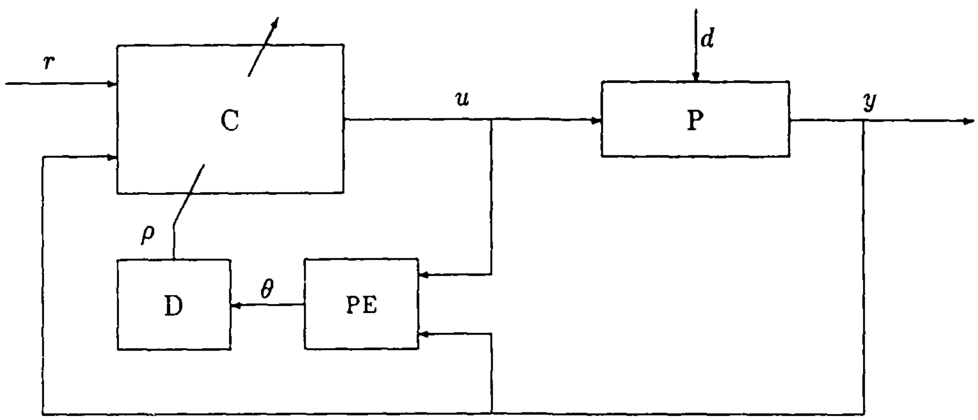

Figure 2.1: Generic Adaptive Control System

The parameter estimator (PE) operates on the input-output data obtained from measurements (y, u) of the plant system (P) and produces a *model parameter estimate* $\theta \in \mathbb{R}^p$. The parameter estimate is transformed by the control design rule (D) into a controller parameter $\rho \in \mathbb{R}^\ell$, which is then used in a pre-determined parametric controller structure (C) in feedback with the actual system (P).

In general, it is very easy to construct an adaptive system: just connect a design rule and an estimator together. However, it is very difficult to insure that the resulting adaptive system will provide acceptable performance. This has been the goal of research in this area for over 30 years.

2.1 History of Adaptive Control Approaches

Over the years, the application of adaptive control has not met with a uniform success. A good example is provided by a brief history of the adaptive control technique developed by Whitaker(1959), later referred to as the "M.I.T.-Rule". The basic idea is to adjust some controller parameters in accordance with an instantaneous gradient descent of the squared error between the plant output and the output of a so-called reference model. Some early applications to flight control were reported to be quite successful *e.g.*, [Whitaker(1959)] and [Osborne *et al.* (1961)]. Later reported applications–again to flight control– were not deemed successful, *e.g.*, [Bryson(1977)]. At the same time as these mixed results were being obtained in the area of flight control applications, several researchers were reporting good results with applications in process control, engine control, and ship steering, *e.g.*, [Draper and Li(1951), Astrom *et al.* (1965,1973)]. Probably because of these mixed results and lack of theory, other approaches to adaptive control were explored.

Analysis via Passivity

One such avenue was announced in a paper by Parks(1966), where the M.I.T.-Rule was "re-designed" in such a way that the resulting adaptively controlled system had a *guaranteed* stability property. The idea is to select the reference model so that a certain (closed-loop) transfer function has a passivity property, specfically, strictly positive real (SPR).[1] With only this SPR information about this transfer function, the adaptive parameters can always be adjusted towards the minimum of some objective function. This "SPR-Rule" idea was developed later for more general situations, *e.g.*, Monopoli(1974), Egardt(1979), and Narendra, Lin, and Valavani(1980). Further extentions and expositions along the SPR line can be found in Landau(1979), Goodwin and Sin(1983), and Kosut and Friedlander(1985).

Effect of Unmodeled Dynamics

Unfortunately, these SPR-Rule algorithms have exactly the same stability problems as the MIT-Rule algorithms, but only when unmodeled dynamics and disturbances are taken into account. Although an early account was given by Egardt(1978), this issue was perhaps not fully realized until the appearance of the work of Rohrs *et al.* (1982), which vividly demonstrated the potentially local unstable nature of adaptive systems by posing a "counter-example" representative of a non-ideal, but practical, circumstance. The adaptive parameters exhibited the characteristic rapid transient followed by a steady parameter drift. In this case, however, the parameters did not settle down in the constant parameter stability set. Thus, once the parameters drifted outside the stability set, the states of the controlled system became exceedingly large.

Cause of Parameter Drift Instability

In scanning the literature, the earliest reference which contains a rigorous analysis of this instability phenomena appears to be in a note by James(1971). A Floquet analysis is applied to an MIT-Rule one-parameter linear adaptive system (adaptive feedforward only) with periodic inputs. The result is a complex stability-instability boundary revealing multiple resonance phenomena, precisely like those associated with the Mathieu equation, *e.g.*, Hale(1969). No analysis was undertaken for the case with adaptive feedback, where the system then is nonlinear.

The work of Ljung(1977) which utilizes a technique of stochastic averaging can perhaps be said to herald the beginning of a nonlinear analysis technique, but its applicability to adaptive systems in non-ideal situations was not fully explored at that time. Lyapunov techniques, explored by Anderson and Johnstone(1983) and Ioannou and Kokotovic(1983), showed that *persistently exciting* signals are required to provide a uniform asymptotic stability of the ideal adaptive system, and this in turn provides robustness to various unmodeled dynamics and disturbances. Input-output formulations providing a local stability were developed in Kosut and Johnson(1984) and Kosut and Anderson(1986); these also needed persistent excitation. In none of these examinations of the stability properties of adaptive control systems was the *precise* mechanism for the parameter drift instability identified.

---

[1] Roughly, a transfer function $G(s)$ is SPR if it is stable, and for some constant $\epsilon > 0$, $\mathrm{Re}[G(j\omega - \epsilon)] \geq 0$, $\forall \omega$.

Method of Averaging

The first insights came in a series of papers by Åström (1983,1984), which provided an analysis based on the classical method of averaging of Bogoliubov and Mitropolski(1961). The method requires slow adaptation and/or small signal magnitudes, and under these conditions Åström was able to identify the source of the slow drift instability mechanism. This was made precise in a paper by Riedle and Kokotovic(1984,1985), where they established an *average SPR condition* which provides a sharp stability- instability boundary, again in the case of slow adaptation with periodic inputs, but only for the linearized adaptive system.[2] Extentions of this important result to other than periodic signals and the relation to persistent excitation and unmodeled dynamics is provided in Kosut, Anderson, and Mareels(1986). Extending the method of averaging to the full adaptive feedback case has also been developed, *e.g.*, Riedle and Kokotovic(1986), Kosut(1986,1987), and Bodsen *et al.* (1985).

During this period many of the researchers involved worked more or less in consort. A summary of these early efforts is contained in *Stability of Adaptive Systems: Passivity and Averaging Analysis* (MIT Press) by Anderson *et al.* (1986). The material in the text represents some modifications and refinements of the earlier work mentioned above.

Present Status

In summary, we now understand that the primary cause of poor performance in adaptive systems is that the true plant system is *not* in the parametric model set upon which the parameter estimator is designed. At best, the underlying model set is an approximation, perhaps only valid over a limited range of dynamics. Hence, the states of the parameter estimator cannot be separated from the states of the plant system. In consequence, adaptive algorithms derived from the passivity based theory, which requires this strict separation, can be very sensitive to even slight unmodeled disturbances or high frequency dynamics.

At present, there are basically two research "schools", namely: (1) adapt slowly and insure robustness, and (2) develop adaptive algorithms which are inherently robust. In our research under AFOSR funding we have concentrated on both of these. To accomplish (1) we have applied and extended the method of averaging analysis which leads to the frequency domain average SPR condition mentioned before. To accomplish (2) we have developed what we call "uncertainty estimation" [*cf.* Kosut(1987b)]. The idea is to replace the model estimator in Figure 2.1 with one that produces an estimate together with a measure of uncertainty. The control design rule is replaced with a rule which accepts a set of uncertainty for design. These approaches to (1) and (2) have been developed for linear systems in general and space structures in particular. As part of our future research, we plan to extend these techniques to *nonlinear* flexible systems.

In the next few sections we will show how we plan to develop these methods to further our understanding in the case of nonlinear flexible systems. We first dicsuss the method of averaging in Section 2.2 and then uncertainty estimation in Section 2.3.

[2] The average SPR condition is significantly less restrictive than the usual SPR condition because the latter requires that a certain closed-loop transfer function be positive at *all* frequencies, whereas the former depends also on a signal spectrum and requires that the energy at those frequencies where the transfer is positive should dominate those frequencies where it is negative.

2.2 Averaging Analysis of Adaptive Control Systems

In this section we review the classical method of averaging, review the current status with respect to adaptive control of linear systems, and indicate how we plan to extend the ideas to nonlinear systems. We also provide a motivating example simulation of an adaptively controlled nonlinear system exhibiting chaos.

2.2.1 Classical Method of Averaging

The classical method of averaging analysis considers systems of the form

$$\dot{x} = \gamma f(t, x) \tag{2.1}$$

where γ is a positive constant. Under suitable smoothness conditions the solutions of the above system, for all small $\gamma > 0$, are approximated to order-γ by solutions of the "averaged system"

$$\dot{y} = \gamma \bar{f}(y) \tag{2.2}$$

where $\bar{f}(y)$ is the average over t of $f(t, y)$ for fixed y, assuming the average exists. That is,

$$\bar{f}(y) = \lim_{T \to \infty} \frac{1}{T} \int_0^T f(t, y)\, dt \tag{2.3}$$

The details on averaging theory including the smoothness conditions, averageable functions, *etc.*, are contained in many textbooks, *e.g.*, Bogoliuboff and Mitropolskii (1961), Hale(1969), Sanders and Verhulst(1987). The usefulness of averaging is that the averaged system, although still nonlinear, is simpler to analyze. For example, let y_* denote an equilibrium of the averaged system, that is, y_* satisfies

$$\bar{f}(y_*) = 0$$

If y_* is a stable equilibrium of the averaged system, then there is a stable solution $x_*(t)$ of the original system (2.1) which is contained in an order- γ neighborhood or orbit of y_*. Thus, the original system inherits the stability of the averaged system, which is easily determined from a linearization about the equilibrium y_*. In addition, the region of attraction of the averaged system to the equilibrium y_* is to within order -γ of the region of attraction of the original system to the stable orbit $x_*(t)$.

2.2.2 Current Status: Slowly Adapting Linear Systems

The one difficulty in directly applying the classical method of averaging to adaptive systems is that they do not fit the standard form of (2.1). Specifically, the system of Figure 2.1, as well as most adaptive control systems, can be represented by the set of coupled ordinary differential equations: [3]

$$\begin{aligned} \dot{x} &= g(t, x, \theta) && (2.4) \\ \dot{\theta} &= \gamma q(t, x, \theta) && (2.5) \end{aligned}$$

[3] Since dynamical systems are controlled by a digital computer, by rights we should use difference equations to describe the parameter adaptation. Nonetheless, we use (2.5) to illustrate the main ideas, which for the most part carry forward to the discrete-time case [*cf.* Ch. 5, Anderson *et al.* (1986)].

In (2.4), $x \in \mathbb{R}^n$ is referred to as the "state" and consists of dynamical states of the system being controlled, controller states, and filter states in the parameter estimator. The function $g(t, x, \theta)$ is determined by the system dynamics and the controller/estimator design. In (2.5), $\theta \in \mathbb{R}^p$ is the "parameter estimate" whose rate of change is governed by a scalar constant $\gamma > 0$ and a function $q(t, x, \theta)$ determined by the designer. Observe that when the parameters are held fixed, the system is governed by the nonlinear system (2.4) for constant parameter θ.

In the case of an *adaptive linear system*, (2.4)-(2.5) reduce to,

$$\begin{aligned} \dot{x} &= A(\theta)x + B(\theta)w(t) && (2.6) \\ \dot{\theta} &= \gamma q(t, x, \theta) && (2.7) \end{aligned}$$

where $w(t) \in \mathbb{R}^m$ consists of all exogenous inputs such as references and disturbances, and $A(\theta) \in \mathbb{R}^{n \times n}$ and $B(\theta) \in \mathbb{R}^{n \times m}$ are matrix functions of the parameter θ, the specific function depending on the control design rule and the parameter estimation model set. In this case when the parameters are held fixed, the system is governed by the LTI system (2.6) with θ constant.

Following Anderson *et al.* (1986), to apply the method of averaging to the adaptive linear system (2.6)-(2.7) , we make the assumption that the adaptation gain, γ, is small. This assumption corresponds to slow adaptation, *i.e.*, $\theta(t)$ is slowly changing. To further facilitate the analysis a time scale decomposition follows giving rise to new differential equations that are coupled in a simpler way. First, freeze $\theta(t)$ at a constant value θ, and let $\bar{x}(t, \theta)$ denote the solution to the linear, constant system

$$\left.\begin{aligned} \dot{\bar{x}} &= A(\theta)\bar{x} + B(\theta)w(t) \\ \dot{\theta} &= 0 \end{aligned}\right. \qquad (2.8)$$

called the "frozen system". Observe that $\bar{x}(t, \theta)$ is a function of any vector θ and a scalar t. Hence, we may define the error state

$$\eta(t) = x(t) - \bar{x}(t, \theta(t)) \qquad (2.9)$$

The adaptive linear system (2.6)-(2.7) is then equivalent to

$$\begin{aligned} \dot{\theta} &= \gamma q(t, \bar{x}(t, \theta) + \eta, \theta) && (2.10) \\ \dot{\eta} &= A(\theta)\eta - \gamma \frac{\partial \bar{x}}{\partial \theta}(t, \theta) q(t, \bar{x}(t, \theta) + \eta, \theta) && (2.11) \end{aligned}$$

It is convenient to think of η as the "fast" states, particularly when γ is small. Thus, (2.10)-(2.11) is a time-scale decomposition of (2.6)-(2.7) . Now, if the fast states asymptotically decay, then $\theta(t)$ approaches solutions of

$$\dot{\theta}_m = \gamma q(t, \bar{x}(t, \theta_m), \theta_m) \qquad (2.12)$$

which is in precisely the form of (2.1), and hence, suitable for averaging. Thus, the "averaged parameter system" is

$$\dot{\theta}_a = \gamma f(\theta_a) \qquad (2.13)$$

where

$$f(\theta) = \lim_{T \to \infty} \frac{1}{T} \int_0^T q(t, \bar{x}(t, \theta), \theta)\, dt \qquad (2.14)$$

The important results are roughly as follows:

(i) The parameter trajectories of the adaptive linear system (2.6)-(2.7) are to within an order-γ of the solutions of the averaged system (2.13).

(ii) The stability of the adaptive linear system (2.6)-(2.7) is inherited from the stability of the averaged system (2.13).

(iii) A stable equilibrium of the averaged system (2.13) has a large region of attraction within the constant parameter stability set associated with the frozen parameter system (2.8), *i.e.*, those fixed θ for which $A(\theta)$ is a stability matrix.

These results have a far reaching impact on quantifying engineering guidelines for adaptive control system design of actual systems. In the first place, they provide a sharp stability-instability boundary. Secondly, although it is not so apparent from this limited exposition, the conditions for stabilty of the equilibrium of the averaged system can be formulated in the frequency domain. It then becomes very clear as to the relation between excitation and system dynamics, *e.g.*, [Anderson *et al.* (1986)],[Kosut *et al.* (1986)].

In light of the averaging results it is now possible to more fully understand the conditions for local stability and perhaps more importantly, the causes and mechanisms for instability. In summary, the method of averaging provides a means to assess the stability and instability properties of slow adaptive linear systems. There are of course some restrictions, the obvious one being the need for slow adaptation. More importantly, how slow is slow? The answer is not well quantified at the present time.

To remove the restrictiveness of slow adaptation will require a better understanding of the transient and bifurcating behavior of the adaptive system. Some preliminary results on the transient behavior is presented in Kosut and Bitmead(1986, 1987). For adaptive linear systems, when the adaptation gain γ is incresed, bifurcation and chaotic phenomena have been observed, *e.g.*, [Mareels and Bitmead(1986,1988)], [Salam and Bai(1988)], [Cyr *et al.* (1983)], [Ydstie(1986)], and [Schoenwald *et al.* (1987)]. In these studies, the underlying system is linear. Thus, *rapid* adaptation can induce bifurcating and chaotic behavior even in linear systems. Even under slow adaptation, such phenomena will appear if the system being controlled is nonlinear, as seen in the example simulation to follow.

2.2.3 A Next Step: Slowly Adapting Nonlinear Systems

As outlined above, the method of averaging so far has been extended and applied only for adaptive *linear* control systems, that is, if the adaptive parameters were held fixed, then the resulting closed-loop system would be linear. This is certainly not the case for flexible mechanical systems which have significant friction and backlash as well as actuator saturation. Also, during rapid slewing maneuvers, the effect of the nonlinear kinematics is apparent. Although the method of averaging can handle nonlinear systems in theory, see, *e.g.*, Hale(1969), the application to adaptive *nonlinear* control remains an open area for basic research.

In principal, some of the linearization type results in Anderson *et al.* (1986) will undoubtedly be applicable, modulo some technical conditions. There is some early work, *e.g.*, Volosov(1962), which may be more directly applicable. However, unlike the adaptive linear system, in the nonlinear case the system may exhibit bifurcations and chaotic behavior even though the adaptation is slow. It is these type of issues, which are intrinsically nonlinear, that appear as the next step in providing a more complete and realistic picture of the use and limitations of parameter adaptive control.

An example is presented in the next section which illustrates some of these issues.

2.2.4 Example: Adaptive Control of a Chaotic System

To provide motivation for studying adaptive control of nonlinear flexible systems we will present an example simulation which vividly illustrates the research issues. Consider the Duffing system:

$$\ddot{x} + k\dot{x} + x^3 = \theta r \tag{2.15}$$

where $k > 0$ is a small damping coefficient, the cubic term, x^3, represents a nonlinear spring, θ is a feedforward control gain, and r is a reference command which is to be followed by the displacement x. Prior knowlege of the damping and nonlinear spring term is assumed unavailable, and hence, the the control parameter θ will be adapted as follows:

$$\dot{\theta} = \gamma r(r - x) \tag{2.16}$$

The constant $\gamma > 0$ as before is used to adjust the speed of adaptation. This adaptation rule is a so-called "gradient" rule (sometimes referred to as the "M.I.T.-Rule", [Whitakar(1959)]) because the rate of adjustment is proportional to the negative gradient, with respect to the parameter θ, of an instantaneous error function. In this case

$$\dot{\theta} \sim -\frac{\partial}{\partial\theta}(x - r)^2 \sim \psi(r - x)$$

where ψ is the instantaneous gradient of x, that is,

$$\psi = \frac{\partial x}{\partial\theta}$$

Thus, for constant θ, ψ satisfies the differential equation

$$\ddot{\psi} + k\dot{\psi} + 3x^2\psi = r$$

Since ψ depends on the unknown damping and cubic nonlinearity in the system (2.15), the pure gradient rule $\dot{\theta} = \gamma\psi(r - x)$ cannot be implemented. Using the (crude but simple) approximation $\psi \approx r$ yields the algorithm of (2.16).

Before studying the adaptive system (2.15)-(2.16) , recall that the Duffing system (2.15) for *constant* θ has been extensively examined. In particular, Ueda(1980) made an exhaustive study with

$$\theta r = B\cos t$$

and tabulated the resulting long-term behavior as a function of the parameters (k, B) as shown in Figure 2.2.

For example, with $(k, B) = (.08, .2)$, there are five coexisting periodic attractors as shown in Figure 2.3. For $(k, B) = (.05, 7.5)$, $(.25, 8.5)$, or $(.1, 12.)$, the attractors are all chaotic, as seen from the $(x, \dot{x})$-plane Poincare' sections at the strobe times $t \in \{2\pi k : k = 1, \ldots, 4250\}$ in Figure 2.4.

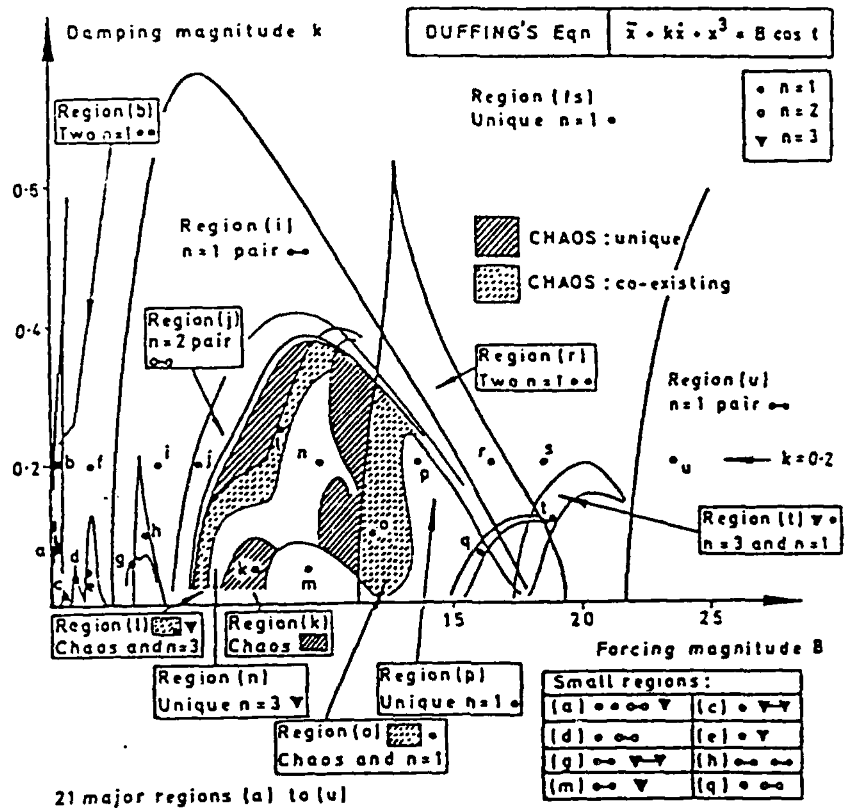

Figure 2.2: Regions of long term behavior of Duffing system in the (k, B) plane. *From Ueda(1980) as reprinted in Thompson and Stewart(1986).*

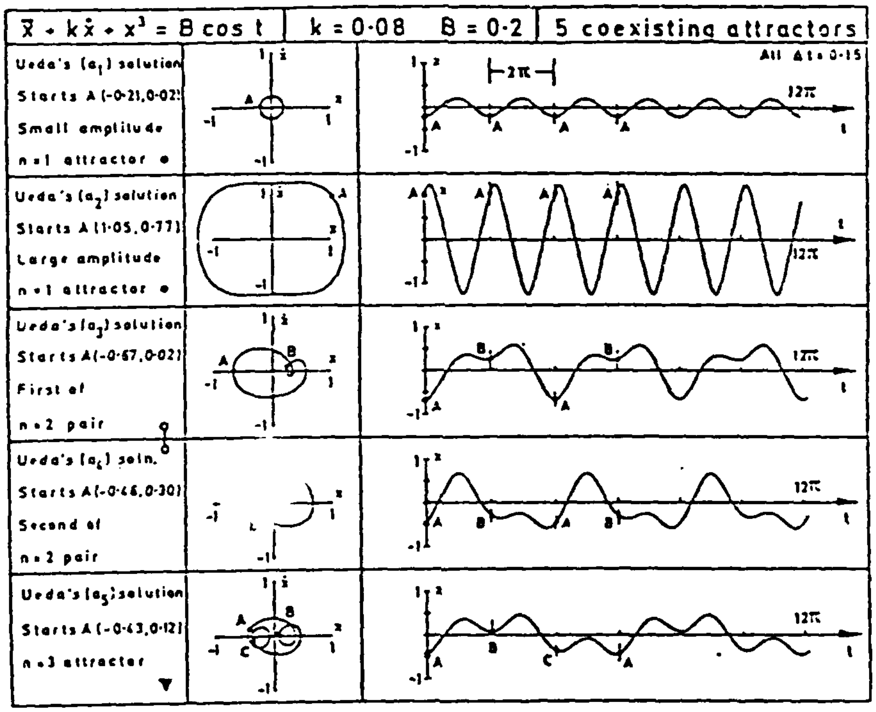

Figure 2.3: Samples of periodic and multi-periodic long-term behavior of the Duffing system under sinusoidal forcing. *From Thompson and Stewart(1986).*

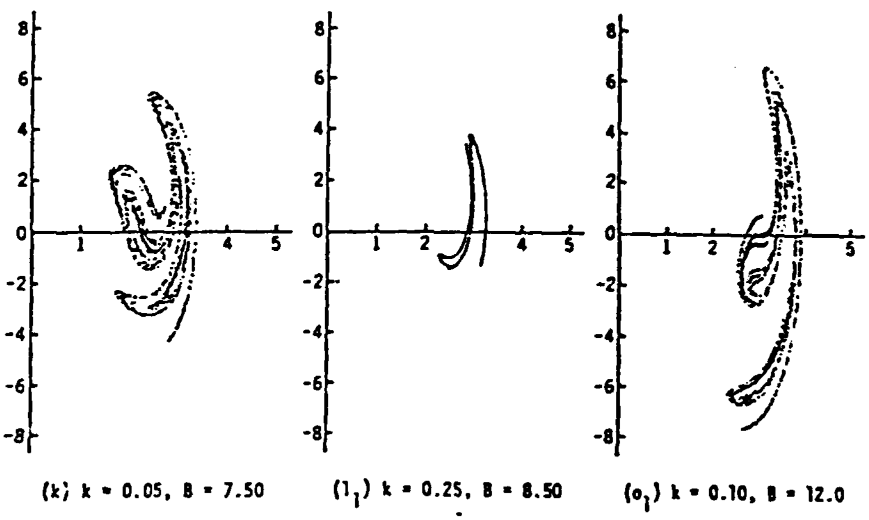

Figure 2.4: Poincare' sections in the $(x, \dot{x})$ plane for three of Ueda's chaotic attractors. *From Ueda(1980) as reprinted in Thompson and Stewart(1986).*

Given the quite complex behavior of the system (2.15) when θ is constant and r is a sinusoidal reference, it is to be expected that if θ is slowly adapted, then the system (2.15)-(2.16) will pass through regions which contain chaotic and/or periodic attractors.

Simulation Results

Simulations of (2.15)-(2.16) were performed under the following conditions:

reference

$$r = .1 \cos t$$

damping

$$k = .05$$

adaptation gain

$$\gamma = \begin{cases} 0 & 0 \le t < 250(2\pi) \\ .01 & 250(2\pi) \le t < 750(2\pi) \\ .05 & 750(2\pi) \le t < 4250(2\pi) \end{cases}$$

initial conditions

$$(\theta, x, \dot{x}) = (75, 0, 0)$$

integration algorithm

4th-order Runge- Kutter integration algorithm with a fixed step-size equal to .005 of the period of the reference signal r, *i.e.*, step-size=(.005)(2π).

To initialize the system there is no adaptation ($\gamma = 0$) for the first 250 periods. This is followed by 500 periods of very slow adaptation ($\gamma = .01$), and then the remainder of the time at relatively slow adaptation ($\gamma = .05$). The adaptive gain, γ, was chosen to produce slow adaptation to prevent any possible bifurcations and chaos soley as a result of too rapid an adaptation, which, as previously mentioned, is known to occur in adaptive systems even when the system being controlled is linear. Thus, initially after the adaptation begins,

$$\theta r \approx 7.5 \cos t$$

Referring to Figure 2.4, the initial response of (2.15)-(2.16) will be chaotic. What we hope to see is that the adaptation brings the system to a less violent condition, which in fact is what occurs. Figures 2.5- 2.7 show the results of the simulation over 4250 periods of the reference, *i.e.*, for $0 \le t \le 4250(2\pi)$.

Figure 2.5 shows values of x and θ at the 1-period strobe times $t \in \{2\pi k : k = 1, \ldots, 4250\}$. Figure 2.6 shows the (x, θ) Poincare' section at these strobe times. Observe that as θ is adapted, the system passes in and out of chaotic and periodic behavior. Specifically, reading Figure 2.5 or 2.6 from left to right: (initially) chaotic $\rightarrow$ 3-periodic $\rightarrow$ chaotic $\rightarrow$ 2-periodic $\rightarrow$ 1-periodic $\rightarrow$ chaotic $\rightarrow$ 3-periodic $\rightarrow$ (finally) 1-periodic. Despite this incredibly complex behavior, the adaptation of θ continues relentlessly to monitonically decrease the parameter value until a good steady state is reached. In fact, some simulations with θ constant (not shown here) verify that the final steady-state adaptive parameter is very near to the constant value of θ minimizing average $(x - r)^2$. This is precisely the desired property of the so-called gradient algorithm (2.16). Figure 2.7 shows some samples of time histories of $(\theta, x, \dot{x})$ vs. t for 5 periods of the reference, and also phase plots of $\dot{x}$ vs. x for the same 5 periods. Comparison of the initial chaotic phase plot with the final phase plot shows a significant improvement in performance in terms of rms error between x and r as well as "mildness" of response.

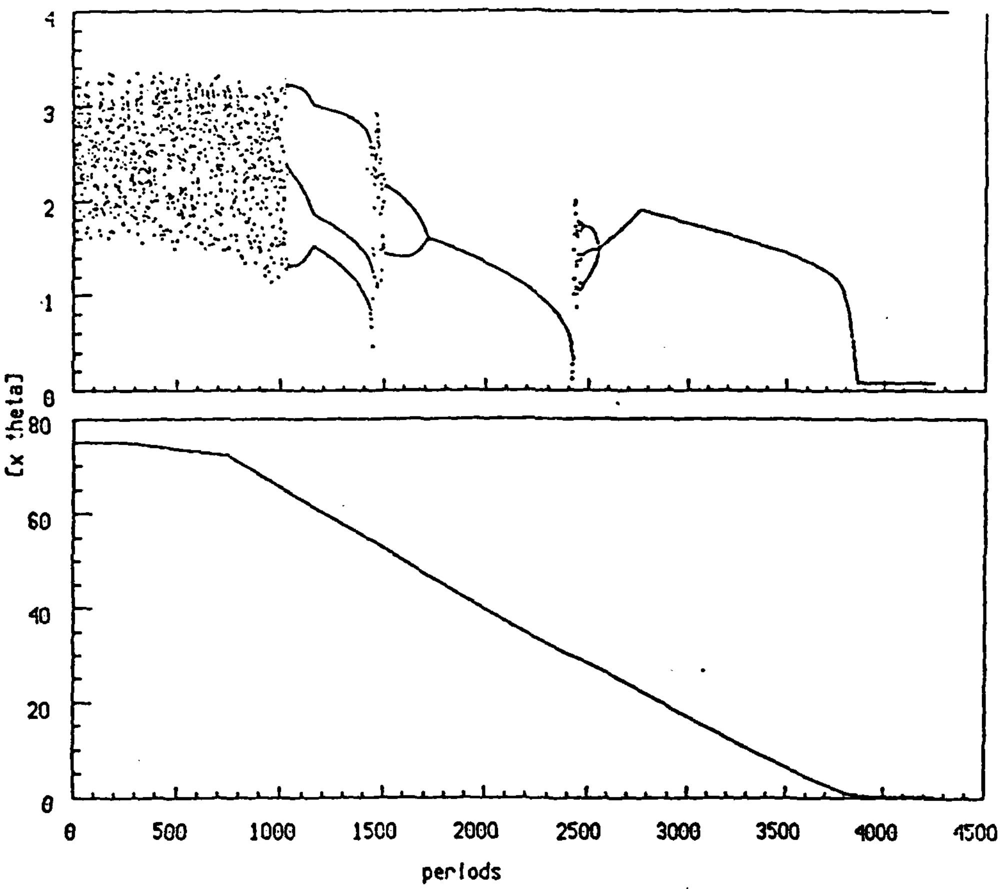

Figure 2.5: Values of x and θ at 1-period strobe times.

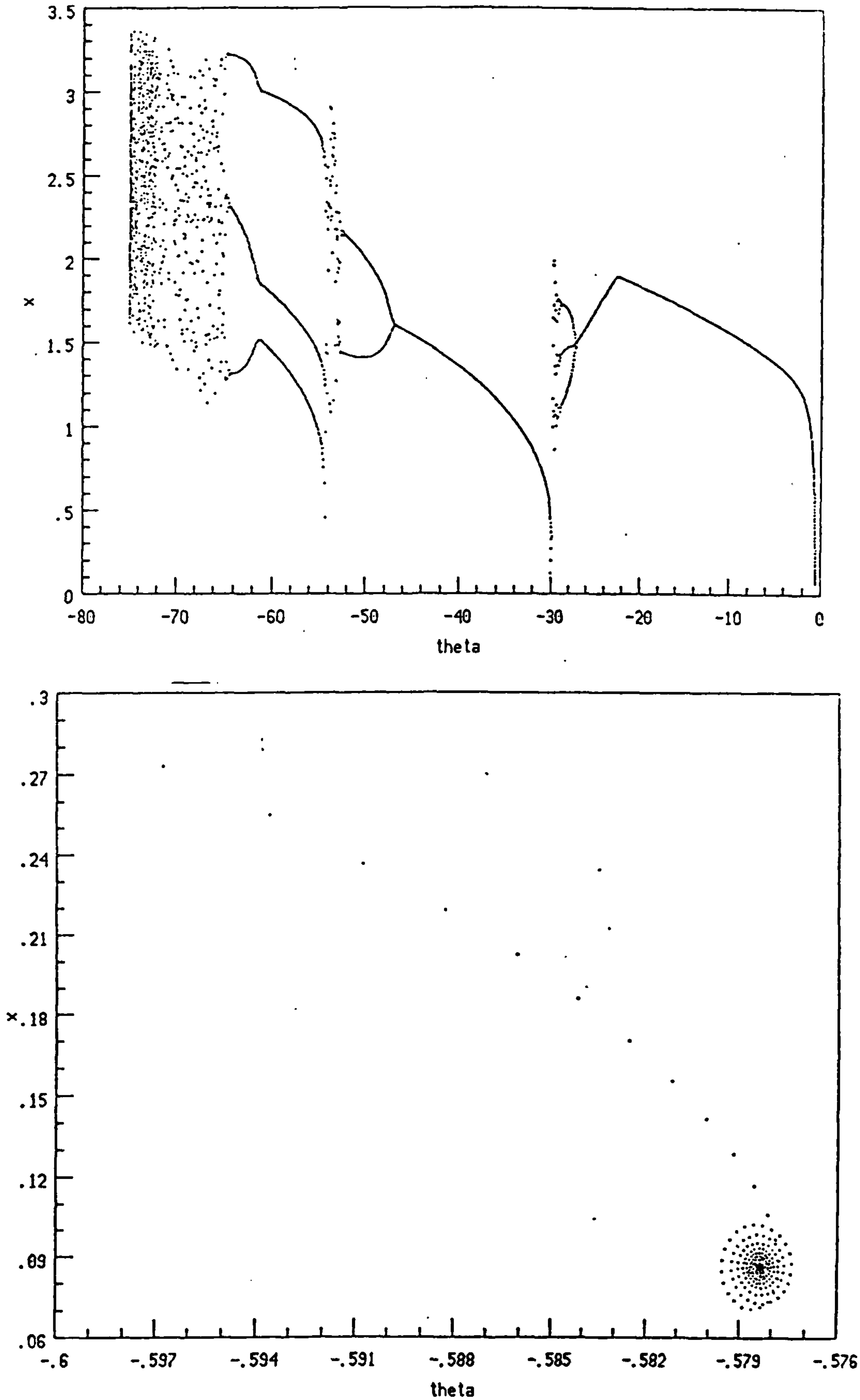

Figure 2.6: *Upper*: Poincare' section in (x, θ). *Lower*: Expanded scale of upper at the end of the simulation.

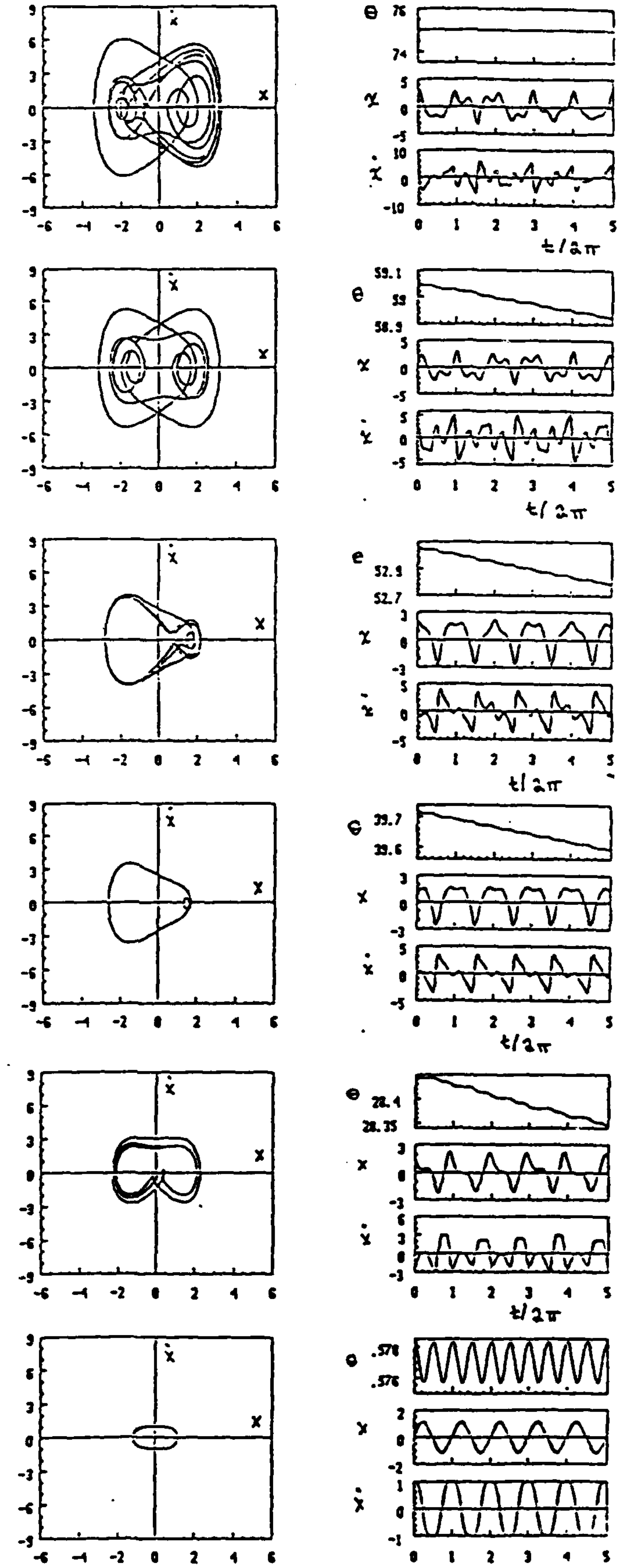

Figure 2.7: Samples of $(\dot{x}, x)$ phase plots and corresponding time histories of $(\theta, x, \dot{x})$ during the simulation.

2.2.5 Implications for Future Research Directions

The simulation results are very encouraging and strongly support the use of adaptive control to minimize the potentially deleterious effects of uncertain nonlinearities. It is also encouraging that the adaptive algorithm (2.16) is very simple to construct and implement. A basic question which arises from the simulation study is simply:

Why does the adaptation work?

In fact, as seen in Figures 2.5 and 2.7, the magnitude of θ decreases monitonically at an almost uniform rate, seemingly independent of whether the system is chaotic or periodic. Moreover, at the end of the simulation, as seen in the lower plot of Figure 2.6, the parameter exponentially spirals towards the final steady-state shown in the bottom time history of Figure 2.7.

At the present time the exact mechanism or conditions which give rise to this pleasing performance are not known. Understanding the underlying principals is a major goal of our research program. Given that slow adaptation is at least prudent, we will certainly consider this situation in our future efforts, and along these lines we can appeal to the method of averaging analysis which we used with success for analyzing adaptive linear control systems under slow adaptation.

2.3 Synthesis via Uncertainty Estimation

In this section we discuss another approach to adaptive control which is referred to as "uncertainty estimation". This approach has been used successfully for linear systems representing LSS in particular, *cf.* Kosut(1987b).

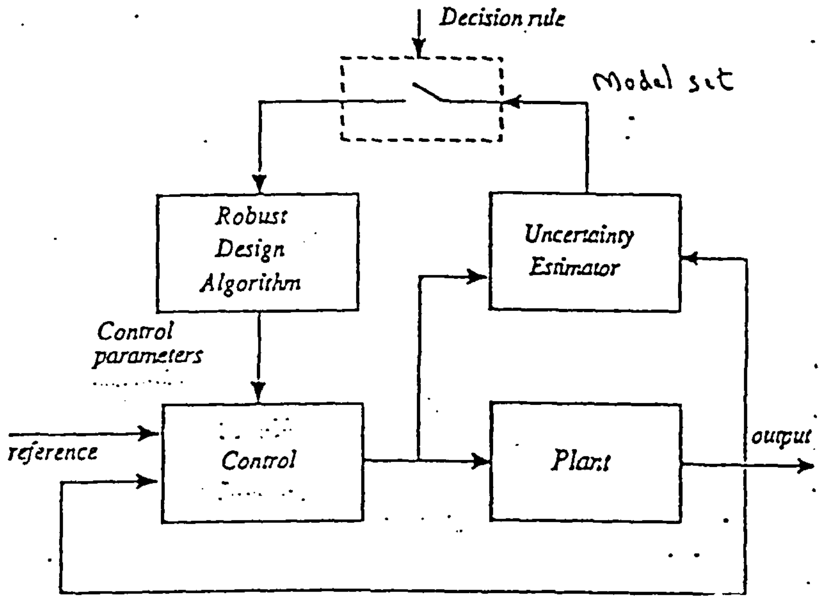

Figure 2.8: Adaptive control with uncertainty estimation.

The uncertainty estimation approach, depicted in Figure 2.8, differs from the usual adaptive system of Figure 2.1 in a number of ways:

- The scheme involves the on-line identification of a *nonlinear model set,* and then tuning or redesigning the controller based on the identified model set.

- The model set generated by the identification process contains information as to the accuracy of the identified model set.

- The control design rule is robust with respect to the identified model set, that is, the closed-loop system satisfies all performance requirements for every plant in the identified model set.

- There is an outer loop containing a "decision rule" as to when to use the model estimate.

This is a fundamentally different approach from the usual adaptive schemes (*cf.* Figure 2.1), where the identification produces a *single* estimated model, with no information regarding its accuracy. In a sense uncertainty estimation shifts the focus from the stability analysis view as expounded in Section 3, to the system identification view.

In this regard there are several basic research issues which need to be resolved for this approach to adaptive control to be effective, namely: (1) model set definition, (2) robust nonlinear control design rules, and (3) uncertainty estimation of nonlinear flexible systems. In the next few subsections we provide a brief description of what is involved in understanding these issues.

2.3.1 Model Set Definition

The starting point in control design or system identification is a definition of a model structure which characterizes the fundamental attributes of the dynamical system. In control design we also need a quantitative description of model accuracy, otherwise there is no way to insure that the design is robust. Thus, the identification process should produce not only a model estimate, but also an estimate of uncertainty. In this task we will define a model set which captures both the dynamical structure as well as being able to describe accuracy.

Specifically, we will begin by examining model sets with the following features:

- *Uncertain Parameters.* A capability to account for that part of the system which is known to be governed by physical laws or able to be described by known functions dependent on certain constant parameters. The parameters may only be known to lie within some range of variation.

- *Uncertain Nonlinearities.* Able to describe uncertainty in the nonlinearities which contribute to the system dynamics. These nonlinearities can still be dependent on uncertain parameters, but here there structure is uncertain as well.

- *Unmodeled Dynamics.* Able to account for uncertain dynamics for which a parametric structure is not available or assumed, *e.g.*, no good physical models or the result of approximations such as neglecting fast dynamics and linearization.

One of our goals is to investigate the use of model sets with the above features for describing uncertain nonlinear flexible systems. In a later section we describe a particular model for the purposes of analysis. However, there are a number of well known methods for describing uncertain nonlinear systems.

Singular Perturbations

In this formulation the sytem dynamics are decomposed into two time scales, where ϵ is a positive constant which reflects the time scale separation between the "slow" states x and the "fast" states z.. This form is known to be compatible with typical areospace systems, see, *e.g.*, Menon *et al.*(1987). However, the singular perturbation parameter ϵ is not easily chosen as a physical parameter, and there is some freedom also in partitioning the states.

The convenience of the above form is that a formal theory exists for separating the time scales, *e.g.*, Chow and Kokotovic(1978).

Conic Sectors

Here we follow in the tradition of Popov(1961), Zames(1966), Safonov (1980,1981), and Doyle(1982,1984), to name a few. The idea of conic-sector modeling is quite general and does encompass the singular perturbation approach described above, however, it is more of an input-output view than a state-space view. The original ideas by Popov and Zames have been generalized by Safonov and Doyle, particularly for LTI sytems. A recent result by Safonov(1987) extends the original sufficient sector stability conditions estiblished by Popov, to necessary conditions, but requires solving a convex programming problem. We hope to explore this latter capability. We will also investigate the optimal scaling methods for LTI conic-sectors established by Doyle(1982,1984).

2.3.2 Robust Nonlinear Control Design

There are now several good methods for designing nonlinear controllers. In this section we discuss a few that can be applied to nonlinear flexible systems.

Feedback Linearization

One particular method which we will investigate is the so called *feedback linearization, e.g.*, Hunt, Su, and Meyer(1983). The basic idea is to apply a nonlinear feedback which brings the closed-loop system into a purely linear form, which has obvious advantages for control design. A general discussion of the theory can be found in Isidori(19?). A difficulty with the method is that the linearizing feedback requires an exact knowledge of the type of system nonlinearities. Robustness to model errors is an open area in basic research. Some work on robustness has been advanced in the robotics area, where feedback linearization is there referred to as "computed torque", see, *e.g.*, Spong and Vidyasagar(1985). Another route is to adapt to maintain parameters in a known nonlinear structure, *e.g.*, Craig, Hsu, and Sastry(1986), Slotine and Li(1987). The more practical adaptive problem would involve parameter adaptation where the true system nonlinearities differ in form from those in the nonlinear model set.

Singular Perturbations

The advantage of the formal time-scale separation discussed before is that the properties of the full nonlinear system can be determined by analyzing the simpler slow and fast nonlinear subsystems. Moreover, the slow and fast controllers can be separately designed.

The time-scale separation parameter ϵ is essentially a measure of the effect of the unmodeled (fast) dynamics on the modeled (slow) part. As mentioned before, it is often a difficult task to define ϵ in terms of physical quantities, and hence, the practicle usefulness involves some guesswork. However, a very nice property of the above approach is that the analysis and formalism is valid for nonlinear systems. Thus, one can start the modeling process with the full nonlinear model, and by introducing ϵ in the right place, a subsequently appropriate model reduction automatically follows. We plan to investigate the applicability of this approach to representing uncertain dynamical systems.

Robust Linear Control

In attempting to develop robust control laws for nonlinear systems we will certainly appeal to existing work on linear systems. For example, robustness of linear feedback systems utilizes frequency domain expressions for characterizing a ***set of uncertainty*** within which lies the true plant, *e.g.*, Francis and Zames(1983), Safonov *et al.*(1981), Doyle and Chu (1985), Vidyasagar(1985). In these approaches, uncertainties in the plant model are described in the frequency domain, and robust stabilization and performance are guaranteed if the peak value, as a function of frequency, of a certain closed loop transfer function is sufficiently small. The peak value turns out to be an appropriate norm in the space of transfer functions, referred to as the H_∞-norm. The design for robust stabilization and performance can then be converted into a minimization problem in the H_∞-norm.

Robust stabilization of linear systems in the presence of *parameter uncertainty* is a rather more difficult problem, and at the present time there is no known solution, although many promising approaches have been advanced. Recently, the analysis and design of robust control systems based on Kharitonov's theorem and its extensions, has received considerable attention in the literature, e.g., see Barmish (1987), and Wei and Yedavalli (1987). Although still in its infancy, this approach will eventually lead to efficient control design methods that can handle structured parameter variations. Methods based on LQG synthesis have also been advanced, *e.g.*, Tahk and Speyer (1986) for general linear systems and Tahk and Speyer(1987) for LSS.

We believe that some of the design methods described in the references above can be further developed by exploring the particular structure of uncertainty occurring in LSS plants. Further, we envision the combination of different design methodologies to achieve the synthesis of controllers that are robust with respect to combined structured and unstructured uncertainty.

2.3.3 Uncertainty Estimation of Nonlinear Systems

Uncertainty estimation is an identification process which produces a model set estimate compatible with prior knowledge regarding uncertainty. The estimator and control design rule use model *sets*, rather than a *single* model estimate as is usually the case. These model sets contain information on model accuracy, *i.e.*, accuracy of parameter estimates and unmodeled dynamics.

There are two types of problems in designing "uncertainty" extimators:

- *Estimation With No Prior Knowledge of Uncertainty*
  In this case there is no knowledge available about the extent of uncertainty in the unmodeled dynamics. The indentification scheme should be modified so as to produce not only an estimated model, but also a measure of model uncertainty.

- *Estimation With Prior Knowledge of Uncertainty*
  In this case a bound on the uncertainty of the unmodeled dynamics is available. The identification scheme should be modified so as to produce a parametric model which is as close to the true system as possible, given the prior information.

2.3.4 Stability Analysis

There is no general theory of stability analysis for adaptive systems which use an uncertainty estimator. In the linear case there are some results based on least-squares estimation, *cf.* Kosut(1988). For nonlinear systems there are some general methods. For example, one can mention the so-called *Total Stability Theorem* for dealing with nonlinear differential equations having the general form

$$\dot{x} = f(t,x) + g(t,x)$$

where $\dot{x} = f(t,x)$ is the "nominal" system and where $g(t,x)$ represents "perturbations." The original form of the theorem is due to Malkin(1958). Later versions can be found in Bellman(1953), Hale(1969), and more recently in Anderson *et al.* (1986). Roughly, the *Total Stability Theorem* asserts that if the nominal system is stable, and the perturbations are bounded and sufficiently smooth, then the perturbed system is also stable. One often tries to make the nomial system linear or of a special nonlinear structure, *e.g.*, adaptive in the ideal case of no unmodled dynamics. This is certainly an avenue which we will explore.

Another route is based on the structure of Figure 2.8. That is, the decision rule switch is open for a period of time while identification takes place. That is, the parameter estimation process is sparate focibly from the control design process. Thus, the control parameters are held fixed at there previous values for a period of time while new values are learned. For example, let θ_{k-1} denote a previously determined parameter value which is being held fixed during the current observation or learning interval $t \in [(k-1)T, kT]$. Thus, the adaptive parameters are in effect determined by the *iterative map*

$$\theta_k = \Gamma_k(\theta_k - 1), \quad k = 1, 2, \ldots$$

where $\Gamma_k(\cdot)$ is defined implicitly by

$$\begin{aligned} \dot{x} &= g(t, x, \theta_{k-1}) \\ \dot{\theta} &= \gamma q(t, x, \theta) \qquad \theta((k-1)T) = \theta_{k-1} \\ \theta_k &= \theta(kT) \end{aligned}$$

An interesting research question is what happens if the observation intervals is long, that is, γ small and T large. From the previous averaging results one would suspect that for all small $\gamma > 0$:

$$\lim_{T\to\infty} \lim_{k\to\infty} \theta_k = \theta_* + \mathcal{O}(\gamma)$$

where θ_* is a *fixed-point* of the map $\Gamma(\theta)$ defined implicitly as the equilibrium of the averaged system in the following sense:

$$\lim_{T\to\infty} \frac{1}{T} \int_{t_0}^{t_0+T} q\,[t, \bar{x}(t, \theta_*), \Gamma(\theta_*)]\, dt = 0$$

Some preliminary results in Phillips, Kosut, and Franklin(1988) suggest that $\Gamma(\cdot)$ is *contractive* in the constant parameter stability set if and only if θ_* is a uas equilibrium of the averaged system (2.13). This conjecture would also tie together some existing results on bounding the actual fixed-point, see, *e.g.*, Kosut(1988).

2.4 Representaive Nonlinear Model and Testbed

In this section we discuss in some detail the type of nonlinear flexible system to be adaptively controlled. In this regard we first discuss a project at ISI which can provide a realistic model base as well as an experimental apparatus.

2.4.1 Adaptive System Testbed

Under an Army funded program at ISI, we are developing an "adaptive system testbed" for rapid prototyping of advanced adaptive control algorithms. The objective of the testbed facility is to be able to design, develop, test and validate weapon system control strategies so that a high degree of confidence can be established in transferring the technology to test vehicles and operational systems. The testbed is specifically designed for testing weapon pointing control of flexible systems.

Hardware Configuration

The proposed system for the rapid prototyping of control algorithms will consist of an Apollo workstation, a multi-processor real-time controller and the testbed (see figure 2.9). The Apollo workstation will be used for developing and simulating the control algorithms, for generating real-time code, and for interfacing with the real-time control processor. The real-time control processor *will execute the code* generated for the control algorithm and interface with the testbed.

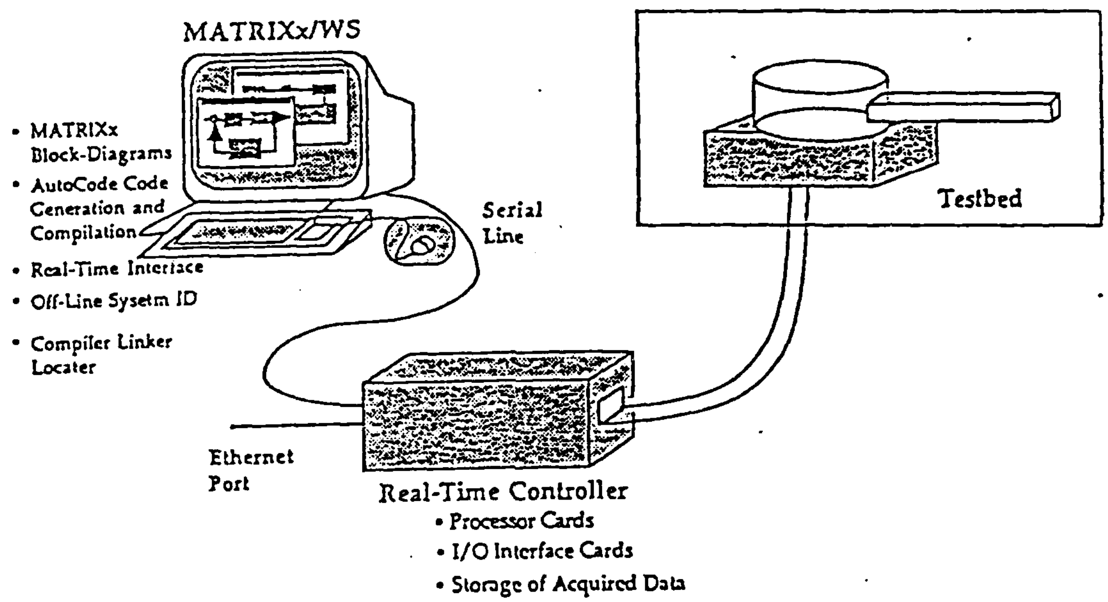

Figure 2.9: Adaptive Control Development System

Software Elements

The feature making the Adaptive Control testbed most desirable for control system experimentation is the ability to easily and rapidly reprogram the processors. Controller architectures and parameters are specified in block diagram form at a high level using SYSTEM_BUILD, then the SYSTEM_BUILD diagrams are converted to real-time control programs through an automatic code generator. This code is then compiled and downloaded to all or some of the processors for real-time execution. The user can interface with the code executing on the controller through graphical icons on the host workstation.

Mechanical Testbed Fixture

A test fixture (see Figure 2.10) is proposed so that variations in crucial physical parameters can be quantified and tested easily, for example: drive compliance, drive freeplay (backlash), barrel length, disturbances, and sensors. Physical variations will be accomplished where possible by interchangeable and mechanically adjustable elements, otherwise the effects will be emulated in the real time processor.

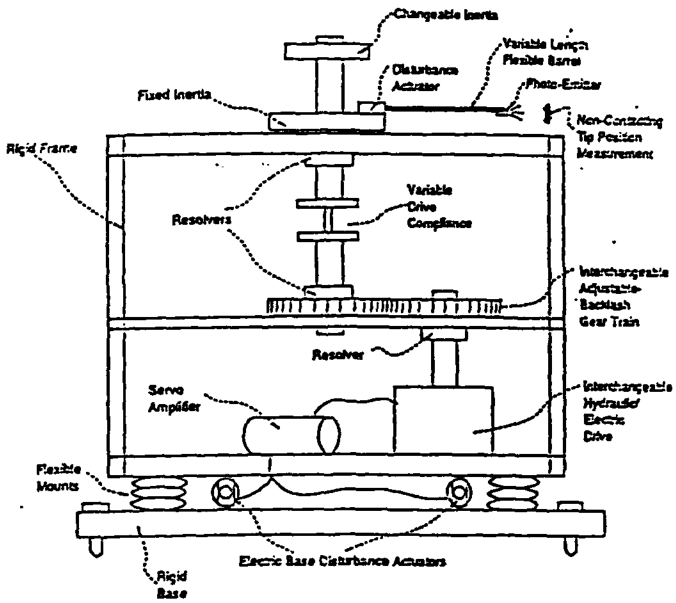

Figure 2.10: Pointing Testbed

By making all of the above functions readily adjustable either mechanically or by emulation in the processors, the performance of different control algorithms based on various sensor and resolver configurations can be compared.

2.4.2 Modeling of Nonlinear Flexible Systems

Although the Duffing system (2.15) used in the simulation example is interesting, it does not capture all of the essential characteristics of flexible nonlinear systems under control, and moreover, does not well represent the Adaptive Testbed. However, because it has been so extensively studied, we do plan to use it as a stepping stone to more representative systems. In particular, a more representaive scenario is depicted below in Figure 2.11.

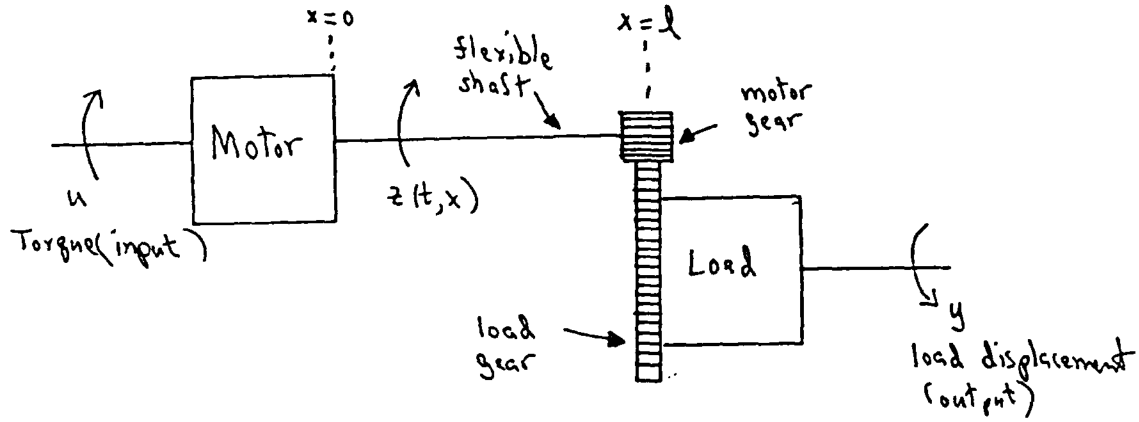

Figure 2.11: A flexible rotating system with backlash in the gear-train.

This system represents the case where the actuation is applied to a load through a flexible and possibly geared device. This is a common situation in mechanical systems such as robots or in proposed active controllers for large space structures. The gearing is shown to occur at the end of the flexible member, although other combinations may be possible, this perhaps being the more common.

Neglecting any electronic dynamics, and assuming that the flexible rod is both uniform and undamped, the motion of the system for small angular deflections is well approximated by the partial differential equation in time t and spatial coordinate x,

$$\mu z_{tt}(t,x) - \rho z_{xx}(t,x) = 0 \tag{2.17}$$

with boundary conditions

$$\begin{aligned} \rho z_x(t,0) &= J_M z_{tt}(t,0) - u(t) & (2.18)\\ \rho z_x(t,\ell) &= -J_G z_{tt}(t,\ell) + g[\alpha(t)] & (2.19)\\ J_L \ddot{y}(t) &= N g[\alpha(t)] & (2.20) \end{aligned}$$

where y is the load output, u is the user applied torque, and α is the relative gear angle

$$\alpha(t) = z(t,\ell) - Ny(t) \tag{2.21}$$

The constants are defined as follows: J_M, J_G, and J_L are the motor, motor gear, and load inertias, respectively, N is the gear ratio which is greator than one, μ is the tortional mass density, and ρ is the tortional stiffness. The nonlinear function $g(\cdot)$ arises from backlash in the gear train, and has the typical shape as shown in Figure 2.12.

The break-point parameter α_b relates to gear teeth spacing and the slopes in the two regions relate to gear teeth shapes. Typically for $|\alpha| > \alpha_b$ the slope is very large whereas for $|\alpha| < \alpha_b$ the slope is very small.

Observe that (2.17)-(2.20) contains many of the features of a realistic flexible system. Firstly, it is a PDE and thus captures the large number of modes which characterize flexible systems. Secondly, the nonlinearity $g(\cdot)$ is common to practically all actuators, both in robotics and those proposed for space structure active control. Moreover, this type of nonlinearity is also representative of the interaction between joints and flexible structural members in a space structure. Of course

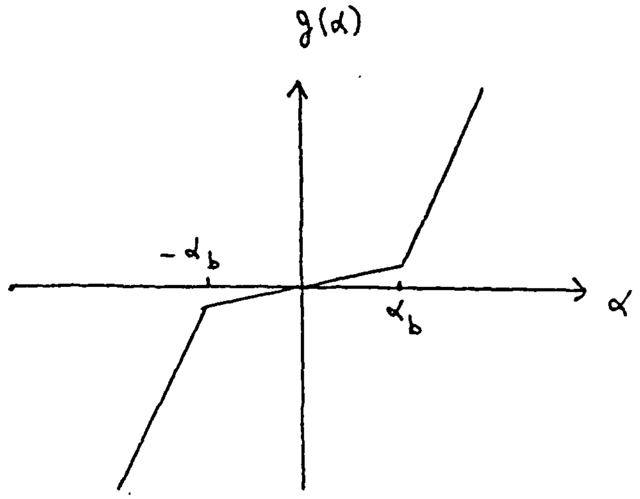

Figure 2.12: A typical backlash function.

there is no real end to posing even more complicated models, and at some level even (2.17)-(2.20) can be accused of being too simple. The same applies to the Duffing system (2.15), but the counter argument is based on the rich complexity which results from this "simple" model, *e.g.*, Figure 2.2. Nonetheless, we take the position that (2.17)-(2.20) is our representaive system for the purposes of analysis. As mentioned before, we still plan to use the adaptive Duffing system (2.15)-(2.16) as a stepping stone to further our understanding of systems like (2.17)-(2.20) .

As an intermediate step we will also make use of the following ODE as a one-mode approximation of (2.17)-(2.20) which includes damping:

$$J_M \ddot{z}_1 = u - D(\dot{z}_2 - \dot{z}_1) + K(z_2 - z_1) \tag{2.22}$$
$$J_G \ddot{z}_2 = -g(\alpha) \tag{2.23}$$
$$J_L \ddot{y} = N g(\alpha) \tag{2.24}$$
$$\alpha = z_2 - N y \tag{2.25}$$

where z_1, z_2 are defined by

$$z_1(t) = z(t,0) \qquad z_2(t) = z(t,\ell) \tag{2.26}$$

and where the constants D and K represent damping and stiffness, respectively.

2.5 References

B.D.O. Anderson, R.R. Bitmead, C.R. Johnson, Jr., P.V. Kokotovic, R.L. Kosut, I.M.Y. Mareels, L. Praly, and B.D. Riedle, (1986),*Stability of Adaptive Systems: Passivity and Averaging Analysis*, MIT Press, 1986.

B.D.O. Anderson and R.M. Johnstone (1983), "Adaptive Systems and Time- Varying Plants", *Int. J. Control*, Vol. 37, No. 2, pp.367-377.

K.J.Åström (1983),"Analysis of Rohrs Counter-Examples to Adaptive Control", *Proc. 22nd IEEE Conf. on Dec. and Contr.*, San Antonio, TX.

K.J.Åström (1984),"Interactions Between Excitation and Unmodeled Dynamics in Adaptive Control",*Proc. 23rd IEEE Conf. on Dec. and Contr.*,Las Vegas, NV, pp.1276-1281.

K.J. Åström and P. Eykhoff (1971), "System Identification—A survey", *Automatica*, 7:123–167.

N.N. Bogoliuboff and Y.A. Mitropolskii (1961) *Asymptotic methods in the theory of Nonlinear Oscillators*, Gordon and Breach, New York.

M. Bodson, S. Sastry, B.D.O. Anderson, I. Mareels, and R.R. Bitmead, (1986), "Nonlinear Averaging Theorems, and the Determination of Parameter Convergence Rates in Adaptive Control", *Systems and Contr. Letters*, to appear.

A. Bryson (1977), Guest Editorial in Mini-Issue on NASA's Advanced Control Law Program for the F-8 DFBW Aircraft, *IEEE Trans. Aut. Control*, Vol. AC-22, No.5, Oct. 1977.

M. Bodson, S. Sastry, B.D.O. Anderson, I. Mareels, and R.R. Bitmead, (1986), "Nonlinear Averaging Theorems, and the Determination of Parameter Convergence Rates in Adaptive Control", *Systems and Contr. Letters*, to appear.

S. Boyd et al. (1986), "A New CAD Method and Associated Architectures for Linear Controllers", ISL Tech. Report L-104-81,1, Stanford University, Dec. 1986.

D.R. Brillinger (1975), *Time Series: Data Analysis and Theory*, Holt, Rinehart, and Wintron, New York.

B. Cyr, B.D. Riedle, and P.V.Kokotovic (1983), "Hopf bifurcation in an adaptive system with unmodelled dynamics," *Proc. IFAC Workshop on Adaptive Systems*, San Francisco, CA.

C.A. Desoer, R.W. Liu, J. Murray and R. Sacks (1980), "Feedback System Design: The Fractional Representation Approach to Analysis and Synthesis", *IEEE Trans. Aut. Control*, Vol. AC-25, pp. 399–412, June 1980.

C.A. Desoer and M. Vidyasagar, *Feedback Systems: Input-Output Properties*, Academic Press, 1975.

J.C. Doyle and G. Stein (1981), "Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis," *IEEE Trans. Automat. Contr.*, Feb. 1981.

J.C. Doyle and C.C. Chu (1985), "Robust Control of Multivariable and Large Scale Systems", *Honeywell SRC: Final Technical Report* on AFOSR Contract F49620-84-C-0088, March 1986.

C.S. Draper and Y.T. Li (1951), *Principals of Optimalizing Control Systems and an Application to the Internal Combustion Engine*, American Society of Mechanical Engineers Publication, Sept., 1951.

B. Egardt (1979), *Stability of Adaptive Systems*, Springer-Verlag.

B.A. Francis and G. Zames (1984), "On H_∞-optimal Sensitivity Theory for SISO Feedback Systems," *IEEE Trans. Automat. Contr.*, Vol. 29, pp. 9-16.

G.C. Goodwin and K.S. Sin (1984), *Adaptive Filtering, Prediction, and Control*, Prentice-Hall, New Jersey.

J.K. Hale (1980), *Ordinary Differential Equations*, Kreiger, Molaban, FL, originally published (1969), Wiley (Interscience), New York.

P.A. Ioannou and P.V. Kokotovic (1983) *Adaptive Systems with reduced models*, Springer-Verlag, New york.

D.J.G. James (1971),"Stability of a Model Reference Control System", *AIAA Journal*, Vol. 9, pp.950-952.

G.M. Jenkins and D.G. Watts (1968), *Spectral Analysis and Its Applications,* Holden-Day, San francisco.

R.L. Kosut, B.D.O. Anderson, and I.M.Y. Mareels,(1987), "Stability Theory for Adaptive Systems: Methods of Averaging and Persistency of Excitation", *IEEE Trans. on Aut. Contr.*, Vol. AC-32, No. 1, pp. 20–34, Jan. 1987.

R.L.Kosut, I.M.Y.Mareels, B.D.O.Anderson, R.R.Bitmead, and C.R. Johnson, Jr., (1987),"Transient Analysis of Adaptive Control", *Proc. IFAC, 10th World Congress*, Munich, Germany, July 1987.

R.L. Kosut and R.R. Bitmead (1986), "Fixed Point Theorems for the Stability Analysis of Adaptive Systems", *IFAC Workshop on Adaptive Control*, AC-30:834.

R.L. Kosut and M.G. Lyons, (1986), "Adaptive Control Techniques for Large Space Structures", *Annual Technical Report* (ISI Report No. 85), AFOSR Contract F49620-85-C-0094, Sept. 15, 1986.

R.L. Kosut and B.D.O. Anderson,(1986), "Local Stability Analysis for a Class of Adaptive Systems", *IEEE Trans. Auto. Contr.*, Jan. 1986.

R.L. Kosut and B. Friedlander (1985), "Robust Adaptive Control: Conditions for Global Stability", *IEEE Trans. on Aut. Contr.*, AC-30(7):610–624.

R.L. Kosut and C.R. Johnson, Jr. (1984), "An Input-Output View of Robustness in Adaptive Control", *Automatica*, 20(5)569–581.

R.L. Kosut, (1986), "Adaptive Calibration: An Approach to Uncertainty Modeling and On-Line Robust Control Design", *Proc. 25th IEEE CDC*, Athens, Greece, Dec. 1986.

R.L. Kosut (1987a), "Adaptive Uncertainty Modeling", *Proc. 1987 ACC*, June 10-12, 1987, Minn., MN.

R.L. Kosut (1987b), "Adaptive Control of Large Space Structures: Uncertainty Estimation and Robust Control Calibration", *Large Space Structures: Dynamics and Control*, ed. S.N. Atluri and A.K. Amos, Springer-Verlag, to appear.

L. Ljung, (1987), *System Identification: Theory for the User*, Prentice- Hall, NJ.

L. Ljung and T. Soderstrom (1983), *Theory and Practice of Recursive Identification*, MIT Press.

L. Ljung (1977),"On Positive real Transfer functions and the Convergence of some Recursive Schemes",*IEEE Trans. Auto. Control*, Vol. AC-22, pp.539- 551.

L. Ljung (1985), "On the Estimation of Transfer Functions," *Automatica,* 21(6), Nov. 1985.

I.M.Y. Mareels, B.D.O. Anderson, R.R. Bitmead, M. Bodson, and S. Sastry (1986), "Revisiting the MIT-Rule for Adaptive Control", *Proc. 2nd IFAC Workshop on Adaptive Systems*, Lund, Sweden.

I.M.Y. Mareels and R.R. Bitmead (1986), "Nonlinear dynamics in adaptive control: chatic and periodic stabilization," *Automatica*, vol. 22, pp.641-655.

I.M.Y. Mareels and R.R. Bitmead (1988), "Bifurcation effects in robust adaptive control," IEEE Trans. on Circuits and Systems, vol. 35, no. 7, pp. 835-841, July 1988.

R.V. Monopoli (1974),"Model Reference Adaptive Copntrol with an Augmented Error Signal", *IEEE Trans. Auto. Control*, AC-19, pp.474-484.

K.S. Narendra, Y.H. Lin, and L. Valavani, "Stable Adaptive Control Design, Part II: Proof of Stability", *IEEE Trans. Auto. Control*, Vol. AC-25, No. 3, pp.440-449.

P.V. Osburn, H.P. Whitaker, and A. Kezer (1961),"New Developments in the Design of Model Reference Adaptive control", *Inst. Aeronautical Sciences*, Paper 61-39.

P.C. Parks (1966),"Lyapunov Redesign of Model Reference Adaptive Control Systems", *IEEE Trans. Auto. Control*, Vol. AC-11, pp.362-367.

M.B. Priestly (1981), *Spectral Analysis and Time Series,* Academic Press, New York.

B.D.Riedle and P.V.Kokotovic (1984), "Bifurcating equilibria in a simple adaptive system: simulation evidence," Proc. 1984 ACC, pp.238-240, San Francisco, CA.

B.D. Riedle and P.V. Kokotovic (1986), "Integral Manifolds of Slow Adaptation", *IEEE Trans. Aut. Control*, Vol. 31, No. 4, pp. 316–324, April 1986.

B.D. Riedle and P.V. Kokotovic (1985), "A Stability-Instability Boundary for Disturbance-Free Slow Adaptation and Unmodeled Dynamics", *IEEE Trans. on Aut. Contr.*, AC-30:1027–1030.

C.E.Rohrs, L.S.Valavani, M.Athans, and G.Stein (1985),"Robustness of Continuous Time Adaptive Control Algorithms in the Presence of Unmodeled Dynamics",*IEEE Trans. Aut. Contr.*, AC-30(9):881-889, Sept.

D. Schoenwald, P.V. Kokotovic, S. Dasgupta (1987), "A boundedness conjecture for locally unstable adaptive estimation of non-SPR transfer functions," *Proc. Conf. on Inform. Sciences and Systems*, Baltimore, MD, March.

M.G. Safonov, A.L. Laub, and G.L. Hartmann, (1981),"Feedback Properties of Multivariable Systems: The Role and Use of The Return Difference Matrix", *IEEE Trans. Aut. Contr.*, vol AC-26, Feb. 1981.

F.M.A. Salam and S. Bai (1988), "Complicated dynamics of a prototype continuous-time adaptive control system," *IEEE Trans. on Circuits and Systems*, vol.35, no.7, pp. 842-849, July 1988.

M. Tahk and J.L. Speyer (1987), "Modeling of Parameter Variations and Asymptotic LQG Synthesis," *IEEE Trans. Automat. Contr.*, September 1987.

M.Tahk and J.L. Speyer (1987), " A Parameter Robust LQG Design Synthesis with Applications to Control of Flexible Structures" *Proc. 1987 ACC*, Minneapolis, MN, June 1987.

M. Vidyasagar (1985), *Control System Synthesis: A Factorization Approach*, MIT Press, Cambridge, MA, 1985.

B. Wahlberg and L. Ljung, (1986), "Design Variables for Bias Distribution in Transfer function estimation", *IEEE Trans. Auto. Control*, vol. AC-31, no. 2, pp.134-144, Feb.

H.P. Whitaker (1959),"An Adaptive System for the Control of Aircraft and Spacecraft", *Inst. Aeronautical Sciences*, Paper 59-100.

B.E. Ydstie (1986), "Bifurcations and complex dynamics in adaptive control systems," *Proc. 25th IEEE CDC*, pp.2232-2236, Athens Greece, Dec. 1986.

G. Zames, (1966) "Feedback and Optimal Sensitivity: Model Reference Transformations, Multiplicative Seminorms and Approximate Inverses", *IEEE Trans. Auto. Contr.*, **AC-26**:301–320, April 1981.

Appendix A

Uncertainty Estimation

Adaptive Control via Parameter Set Estimation

ROBERT L. KOSUT*†

Abstract

An adaptive control system is examined where the traditional parameter estimator is replaced by a parameter *set* estimator which provides a measure of uncertainty of the estimated parameters. It is shown how to construct a parameter set estimator with the property that the true system is always in the model set, and furthermore, how to design the estimation experiment so that the set of uncertainty is as small as possible given some *a priori* information.

To appear:
Int. J. of Adaptive Control and Signal Processing
Nov. 1988

---

*Dr. Kosut is a Senior Research Scientist at Integrated Systems, Inc., 2500 Mission College Blvd., Santa Clara, CA 95054, and is a Consulting Professor in the Dept. of Electrical Engineering at Stanford University

†Research supported in part by AFOSR, Directorate of Aerospace Sciences, under Contract F49620-85-C-0094, NSF IUC Grant ECS-8605646, and the NSF U.S-Sweden Cooperative Science Program, Grant INT-8620011

1 Introduction

A parameter adaptive control system, as depicted in Figure 1, consists of two processes which make the system adaptive, namely: (1) a model parameter estimator, and (2) a control design rule.

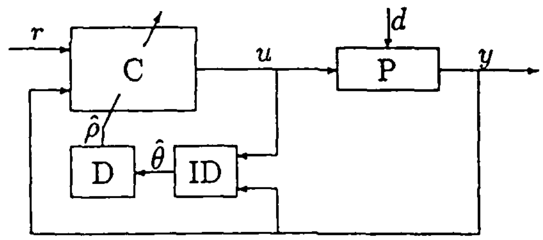

Figure 1: Generic Adaptive Control System

The parameter estimator (ID) operates on the input-output data obtained from measurements (y, u) of the plant system (P) to be controlled, producing a model parameter estimate $\hat{\theta} \in \mathbb{R}^p$. The parameter estimate is transformed by the control design rule (D) into a controller parameter $\hat{\rho} \in \mathbb{R}^\ell$, which is then used in a pre-determined parametric controller structure (C) in feedback with the actual system. Questions regarding the behavior of this system can be answered by describing the estimator properties, in particular, the "goodness" of the "converged" model set.

From the viewpoint of designing a robust control, by a good model set is roughly meant one for which every member produces a controller, via the control design rule, which when applied to the actual system yields acceptable closed- loop performance with regard to attenuating the effect of the disturbance d and tracking the reference r. Hence, a properly designed model estimator should converge to a member of that model set which yields the requsite closed-loop performance.

However, in the adaptive control system of Figure 1, *no specific accuracy information about (P) is passed to the control design rule (D).* Thus, we reach a very important observation:

> *The model estimator should be designed to be compatible with the same class of uncertainty anticipated by the robust control design rule.*

The idea then, is to replace the usual single point parameter estimator of Figure 1 with a *parameter set estimator* which provides a set of parameter values, and hence, a measure of parameter estimation accuracy. The control design rule must then also be modified so as to accept this uncertainty description and produce a controller which is robust to the parameter set.

The estimation of a transfer function from measured data has, of course, a long history, and there will be no attempt to document that here. A good source of references as well as providing a clear elucidation of the technical issues is the textbook by Ljung(1987). One means to provide a measure of uncertainty in the transfer function estimate, is based on utilizing the theoretical asymptotic staitistics to produce confidence intervals from chi-squared distributions. Underlying this approach, however, is the assumption that the true plant is in the model set, and hence, estimation errors are assumed to be dominated by the variance contribution for a sufficiently long observation interval. In the case considered here, the dominant contribution is the bias, which is primarily due to unmodelled dynamics and disturbances. Moreover, it is important when designing the estimation experiment to be aware of the intended use of the estimte, which in the case considered here, is control design.

Some recent work along these lines may be found in Wahlberg and Ljung(1986), Gevers and Ljung(1986), Wittenmark(1987), and Hansen(1988). The idea in these references is to manipulate the user design choices for the identification experiment, such as data filters and input spectra, so that the criterion for estimation looks like the criterion for the intended use of the model estimate, *e.g.,* control design in this case. Another approach is to use high order model sets and then model reduction for control design, *e.g.,* Wahlberg(1986) uses high order least-squares, Parker and Bitmead(1987) use high order FIR models. The appeal of these approaches is that no parametric model need be advanced. In this case the transfer function estimation error consists of a term which depends only on the true transfer function and decays as model order increases, plus a term which depends on the noise-to- signal ratio and may increase as model order increases. These terms can be bounded, but the results are asymptotic giving only qualitative information. Another potential difficulty is the high order which may be too computationally intensive in some situations. The work of LaMaire *et al.* (1987) overcomes the high order problem by using a low order parametric model set. The procedure results in parameter estimates along with estimating a bound on the unmodelled dynamics. In some cases this will also require considerable computation. In Kosut(1986, 1987), a middle road was taken where least-squares is used to estimate the parametric model and a standard spectral estimator is used on the residual output error resulting in a crude bound on the unmodelled dynamics and disturbance spectrum.

In this paper we concentrate on developing a parameter set estimator from the standard least-squares parameter estimator together with some *a priori* information about the the plant system unmodeled dynamics and disturbances. Least-squares parameter estimation is one of the most widely used methods for obtaining parameter estimates. Its popularity arises undoubtedly because of the ease in finding a solution, *i.e.,* the solution is a unique global minimum and has both a closed-form solution, sometimes called batch-least-squares (BLS), and a recusive solution method, or recursive-least-squares (RLS). The problem or disadvantage is that the resulting model estimate obtained via LS will be "biased" as a result of inevitible modeling approximations, most notably, unmodeled disturbances and dynamics. But, the disadvantage can be alleviated by the advantage. In the approach presented here, we exploit the known solution structure, and together with some *a priori* data, produce a parameter set estimator which contains the true system. We also show how to design the estimation experiment so that this set is as small as possible with respect to the *a priori* information. Specifically, we use a transfer function model set which consists of a parametric part with a known structure, and an unstructured part which is known *a priori* to be bounded in the frequency domain. Once this is done, two results follow, namely, (1) an upper bound on *parameter bias* can be determined from the on- line data, and (2) the bound can be made small by appropriate choice of data filters and input spectra. This approach follows quite closely to that of Morrison and Walker(1988), the results here being applicable to more general model sets. Also, the upper bound on parameter error obtained here is asymptotically tight given the *a priori* data, and can be computed using the standard DFT. A preliminary version of this paper is in Kosut(1988).

Other important estimator requirements, such as the rate of convergence and region of attraction, which play a significant role in estimator design, are not specifically addressed. For example, if convergence to the model set is slow, and the initial model estimate is too coarse, then unacceptable behavior may occur during the learning process. Designing adaptive control systems with a prescribed rate of convergence or region of attraction is not a solved problem, although much is understood in the case of slow adaptation, *e.g.*, Anderson et al.(1986).

2 Parametric Transfer Function Modeling

2.1 Sampled-Data Structure

We start with the assumption that the true plant system is accessed through a sampled-data structure where control commands and data acquisition occur at simultaneous discrete sampling instances separated by uniform intervals. Thus, the true plant system can be described by the discrete-time relation

$$y = Gu + d \tag{1}$$

where $y(t)$ and $u(t)$ are the measured output and input sequences, respectively, evaluated at the normalized sample times $t \in \{\ldots, -1, 0, 1, 2, \ldots\}$. The sequence $d(t)$ represents the effect of disturbances sources as seen at the output. The operator G is linear and time-invariant with transfer function $G(q)$, where q is the shift operator, *i.e.*, $(q^k x)(t) = x(t+k)$ for any integer k. Unless needed for clarification we will suppress the 'q' and 't' arguments in transfer functions and sequences, respectively.

Note that the input u may be generated by a feedback mechanism.

2.2 Transfer Function Parametrization

The plant transfer function is assumed to have the following structure:

$$G = \frac{B_{\theta_o} + \Delta_{\mathrm{B}}}{A_{\theta_o} + \Delta_{\mathrm{A}}} \tag{2}$$

where

$$\begin{aligned} B_{\theta_o} &= L_o + \textstyle\sum_{i=1}^{m} b_{o,i} L_i \\ A_{\theta_o} &= M_o + \textstyle\sum_{i=1}^{n} a_{o,i} M_i \\ \theta_o &= (a_{o,1} \cdots a_{o,n}\; b_{o,1} \cdots b_{o,m})^T \in \mathbb{R}^p,\; p = n + m \end{aligned} \tag{3}$$

and $L_i, M_i, \Delta_{\mathrm{B}}, \Delta_{\mathrm{A}}$ are stable transfer functions. We refer to $\theta_o \in \mathbb{R}^p$ as the *true parameter*,

$$G_{\theta_o} = \frac{B_{\theta_o}}{A_{\theta_o}} \tag{4}$$

as the *parametric transfer function*, and to $\Delta_{\mathrm{B}}, \Delta_{\mathrm{A}}$ as *nonparametric* or *unmodeled* dynamics in the numerator and denominator, respectively. The following data is assumed to be available *a priori* :

A Priori Data:

(i) L_i, M_i *are known stable transfer functions.*

(ii) $\Delta_{\mathrm{B}}, \Delta_{\mathrm{A}}$ *are uncertain stable transfer functions satisfying:*

$$\begin{aligned} |\Delta_{\mathrm{B}}(e^{j\omega})| &\leq |W_{\mathrm{B}}(e^{j\omega})|,\ \forall |\omega| \leq \pi \\ |\Delta_{\mathrm{A}}(e^{j\omega})| &\leq |W_{\mathrm{A}}(e^{j\omega})|,\ \forall |\omega| \leq \pi \end{aligned} \tag{5}$$

where W_B, W_A are known stable transfer functions.

(iii) The true parameter θ_o is in the set

$$\Theta_{\rm nom} = \left\{ \theta_o \in \mathbb{R}^p : a_i^{\min} \le a_{o,i} \le a_i^{\max},\ b_i^{\min} \le b_{o,i} \le b_i^{\max} \right\} \tag{6}$$

This model structure and *a priori* assumptions define a *family* of transfer functions. Within this family, or model set, are many of the common transfer function model sets used for robust control design. For example, there is the ubiquitous

$$G = \frac{\sum_{i=1}^{m} b_{o,i} q^{-i}}{1 + \sum_{i=1}^{n} a_{o,i} q^{-i}} \cdot L \tag{7}$$

where L represents nonparametric dynamics. The choices of the model elements in (2) depend upon *a priori* information about L. Suppose, for example, that $\widehat{L}$ is an *a priori* estimate of L, and it is known that the relative error $\widetilde{L} = L/\widehat{L} - 1$ satisfies

$$|\widetilde{L}(e^{j\omega})| \le |W_L(e^{j\omega})|,\ \forall |\omega| \le \pi$$

The corresponding choices in (2) are then:

$$\begin{array}{rcl} B_{\theta_o} & = & \sum_{i=1}^{m} b_{o,i} q^{-i} \widehat{L} \\ A_{\theta_o} & = & 1 + \sum_{i=1}^{n} a_{o,i} q^{-i} \\ \Delta_B & = & \widetilde{L} B_{\theta_o} \\ \Delta_A & = & 0 \end{array} \tag{8}$$

The transfer functions $\{L_i, M_i\}$ in (2) follow from the above definition and are obviously stable. In addition, the unmodeled dynamics are bounded by:

$$\begin{array}{rcl} W_A & = & 0 \\ W_B & = & \left\{ W \text{ stable} : \sup_{\theta_o \in \Theta_o} |W_L(e^{j\omega}) B_{\theta_o}(e^{j\omega})| \le |W(e^{j\omega})| \right\} \end{array} \tag{9}$$

A similar configuration would follow if the unmodeled dynamics involved the absolute error, *i.e.*, where $\widetilde{L} = L - \widehat{L}$.

Another common model form arises from the theory of stable factorization, see, *e.g.*, Vidyasagar(1985). In this case B_{θ_o} and A_{θ_o} represent stable co-prime factors. For example,

$$\begin{array}{rcl} B_{\theta_o} & = & \sum_{i=1}^{m} b_{o,i} \frac{q^{-i}}{D} \\ A_{\theta_o} & = & \frac{1}{D} + \sum_{i=1}^{n} a_{o,i} \frac{q^{-i}}{D} \end{array} \tag{10}$$

where D is a polynomial in q^{-1} such that $1/D$ is stable. In this case, Δ_B and Δ_A are unmodeled dynamics in the stable factors.

In many instances the plant transfer function is bilinear in a particular *physical* parameter. For example, a mass at the tip of a flexible shaft will enter linearly in both the numerator and denominator of the transfer function from the tip position to the torque at the shaft hub, see, *e.g.*, Rovner and Franklin(1987). In general there are always fewer "physical" parameters than "canonical" parameters such as transfer function coefficients. A general bilinear form for a physical transfer function parametrization is (2) with

$$\begin{aligned} B_{\theta_o} &= L_0 + \textstyle\sum_{i=1}^{p} \theta_{o,i} L_i \\ A_{\theta_o} &= M_0 + \textstyle\sum_{i=1}^{p} \theta_{o,i} M_i \\ \theta_o &= (\theta_{o,1} \cdots \theta_{o,p})^T \end{aligned} \tag{11}$$

where now $\theta_o \in \mathbb{R}^p$ is the true *physical parameter* vector. Observe that (11) allows for common parameters in the numerator and denominator of the parametric model, and hence, is a more general form than (3).

2.3 Caveat Emptor

The plant parametrization (2) is limited by the "locations" of the unmodeled dynamics, Δ_B, Δ_A, which represent relative errors in both the numerator and denominator, respectively. If more is known about the source of the unmodeled dynamics, then there is a tendency towards conservatism which limits the control design performance, see, *e.g.*, Doyle *et al.* (1982). Other more sophisticated structures for the puposes of transfer function parameter estimation can be advanced [*e.g.*, Krause and Khargonnekar(1987)], but these will not be pursued here, since for many cases our proposed structure suffices.

3 Least-Squares Parameter Estimation

3.1 Batch-Least-Squares

The batch-least-squares parameter estimator transforms the measured data $\{y, u : t \in [0, N-1]\}$ into the parameter estimate $\widehat{\theta}_N^{\mathrm{LS}} \in \mathbb{R}^p$ by solving:

$$\widehat{\theta}_N^{\mathrm{LS}} = \arg\min_{\theta \in \mathbb{R}^p} \left\langle \varepsilon_\theta^2 \right\rangle_N \tag{12}$$

where ε_θ is a parametric error sequence generated from the measured data, and $\langle \cdot \rangle_N$ denotes averaging over the N-point observation interval, *i.e.*,

$$\langle x \rangle_N = \frac{1}{N} \sum_{t=0}^{N-1} x(t) \tag{13}$$

The parametric structure of ε_θ is motivated by the model structure of (2). Thus, we form the *(filtered) parametric equation error*

$$\varepsilon_\theta = W_{\mathrm{F}}(A_\theta y - B_\theta u) \tag{14}$$

where W_{F} is a stable *data filter*, and corresponding to the form of (2) and (3),

$$\begin{aligned} B_\theta &= L_o + \textstyle\sum_{i=1}^m b_i L_i \\ A_\theta &= M_o + \textstyle\sum_{i=1}^n a_i M_i \\ \theta &= (a_1 \cdots a_n \; b_1 \cdots b_m)^T \end{aligned} \tag{15}$$

The error can be written as an affine function of θ, that is,

$$\varepsilon_\theta = z - \theta^T \phi \tag{16}$$

with

$$\begin{aligned} z &= W_{\mathrm{F}}\,(M_o y - L_o u) \\ \phi &= W_{\mathrm{F}}\,(-M_1 y \cdots M_n y \; L_1 u \cdots L_m u)^T \end{aligned} \tag{17}$$

We refer to $\phi \in \mathbb{R}^p$ as the *regressor*. Clearly, both $z \in \mathbb{R}$ and $\phi \in \mathbb{R}^p$ are obtained by filtering the measured data. Observe that in the more general parametrization of (11),

$$\begin{aligned} z &= W_{\mathrm{F}}(M_o y - L_o u) \\ \phi &= W_{\mathrm{F}}\,(L_1 u - M_1 y \cdots L_p u - M_p y)^T \end{aligned} \tag{18}$$

With the above definitions of z and ϕ, the familiar solution to (12) is

$$\widehat{\theta}_N^{\mathrm{LS}} = \left\langle \phi \phi^T \right\rangle_N^{-1} \langle \phi z \rangle_N \tag{19}$$

provided the indicated inverse exist. The following result is now easily obtained:

Lemma 1 (Linear Regression Model and Parameter Error) *The true system is equivalent to the linear regression model,*

$$z = \theta_o{}^T \phi + \varepsilon_\Delta + \varepsilon_d \tag{20}$$

where

$$\begin{aligned} \varepsilon_\Delta &= W_F(-\Delta_A y + \Delta_B u) && (21) \\ \varepsilon_d &= W_F(A_{\theta_o} + \Delta_A)\, d && (22) \end{aligned}$$

In addition, the parameter error can be decomposed as follows:

$$\widehat{\theta}_N^{LS} - \theta_o = \widetilde{\theta}_N^\Delta + \widetilde{\theta}_N^d \tag{23}$$

where $\widetilde{\theta}_N^\Delta$ is the error contribution due to nonparametric dynamics and $\widetilde{\theta}_N^d$ is the error contribution from the disturbance. Specifically,

$$\begin{aligned} \widetilde{\theta}_N^\Delta &= \left\langle \phi\phi^T \right\rangle_N^{-1} \langle \phi \varepsilon_\Delta \rangle_N && (24) \\ \widetilde{\theta}_N^d &= \left\langle \phi\phi^T \right\rangle_N^{-1} \langle \phi \varepsilon_d \rangle_N && (25) \end{aligned}$$

The expression (20) can be thought of as a "canonical linear regression model" with a parameter θ_o to be estimated from "measurements" $\{z, \phi : t \in [0, N-1]\}$. The sequences $\varepsilon_\Delta, \varepsilon_d$ represent the effect of the nonparametric dynamics and disturbances, respectively. That is, behavior of z which can not be accounted for by the parametric part $\theta_o{}^T\phi$.

The parameter error expressions in Lemma 1 generalize those found in Morrison and Walker(1988). Observe that in the expression for $\widetilde{\theta}_N^\Delta$, the right hand side is a function of the regressor sequence $\{\phi\}$, which is obtained from measurements, and the unknowns, Δ_B, Δ_A, and θ_o. In the expression for $\widetilde{\theta}_N^d$ the only unknown in the right side is the error sequence $\{\varepsilon_d\}$. Thus, any attempt to provide on-line bounds for the parameter error will involve *a priori* data as well as measured data.

3.2 Recursive-Least-Squares

The recursive-least-squares (RLS) algorithm is [see, *e.g.*, Ljung(1987)]:

$$\begin{aligned} \varepsilon(t) &= z(t) - \phi(t)^T \widehat{\theta}_{t-1}^{RLS} \\ \widehat{\theta}_t^{RLS} &= \widehat{\theta}_{t-1}^{RLS} + R(t)^{-1}\phi(t)\varepsilon(t) && (26) \\ R(t) &= R(t-1) + \phi(t)\phi(t)^T \end{aligned}$$

The algorithm starts at $t = 1$ with initial values $\widehat{\theta}_o$ and $R_o = R_o^T > 0$. For $t \geq 1$ the RLS estimate is explicitly given by

$$\widehat{\theta}_t^{RLS} = \left[\frac{1}{t} R_o + \left\langle \phi\phi^T \right\rangle_t \right]^{-1} \left[\frac{1}{t}\widehat{\theta}_o + \langle \phi z \rangle_t \right] \tag{27}$$

Assuming that $\left\langle \phi\phi^T \right\rangle_t^{-1}$ exists for all $t \geq t_o \geq 1$, then for sufficiently large $t \geq t_o$ and/or small R_o, the RLS estimate is approximately the BLS estimate for a data record of length t, *i.e.*,

$$\hat{\theta}_t^{\mathrm{RLS}} \approx \hat{\theta}_t^{\mathrm{LS}} = \left\langle \phi\phi^T \right\rangle_t^{-1} \left\langle \phi z \right\rangle_t \tag{28}$$

A principal adavantage of RLS over BLS is that memory requirements can be vastly different especially for very long data records. BLS stores records of length $\mathcal{O}(t)$, whereas RLS stores only $\mathcal{O}(p)$. Since the RLS estimate will approach the BLS estimate, RLS is usually the method of preference even if the computations are to be done off-line, *e.g.*, RLS is used repeatedly over the data set.

3.3 Parameter Set Estimation

The problem of parameter set estimation is to transform the measured data $\{y, u : t \in [0, N-1]\}$ plus any *a priori* data into the parameter set estimate

$$\hat{\Theta} = \left\{ \theta \in \mathbb{R}^p : \|\theta - \hat{\theta}\| \leq \hat{\delta} \right\} \tag{29}$$

where $\|\cdot\|$ is a vector norm on $\mathbb{R}^p$, $\hat{\theta}$ is a nominal estimate, and $\hat{\delta}$ specifies its uncertainty in $\mathbb{R}^p$. Conceptually, the set estimator is the map

$$\left\{ \begin{array}{c} \{y, u : t \in [0, N-1]\} \\ + \\ a\ priori\ \ \text{data} \end{array} \right\} \mapsto \hat{\Theta} \tag{30}$$

In order for $\hat{\Theta}$ to be useful for control design we first require that the true parameter θ_o is in this set, that is,

$$\theta_o \in \hat{\Theta} \tag{31}$$

Secondly, the set should be sufficiently small so as to insure a robust control design which also satisfies some performance specifications. Hence, a second constraint is

$$\hat{\delta} < \delta_{\mathrm{spec}} \tag{32}$$

where δ_{spec} is obtained from closed-loop design specifications. Without the design constraint the estimation problem has no practical meaning, because then typically the *a priori* set is sufficient for design; the result being a stable, but low performance closed-loop system.

4 Parameter Error From Unmodeled Dynamics

Using Lemma 1, we will establish a bound on $\|\tilde{\theta}_N^{\Delta}\|$ which depends on the measured and *a priori* data. The norm used is the ℓ_∞-norm on $\mathbb{R}^p$, that is, $\|x\| = \max\{|x_i| : i \in [1,p]\}$, which is consistent with independent model parameters as in (6). One can rationalize other choices as well. For example, the ℓ_2-norm, $\|x\| = (x^T x)^{1/2}$, implicitly imposes a relation amongst the parameters which may be sensible in some problem context.

In the results to follow we use $\|\cdot\|$ to also denote the matrix norm induced by the ℓ_∞-norm, that is, for any matrix A with (possibly complex) elements a_{ij},

$$\|A\| \stackrel{\text{def}}{=} \sup_{\|x\|=1} \|Ax\| = \sup_i \sum_j |a_{ij}| \tag{33}$$

We also use 'abs$[A]$' to denote the matrix whose elements are the absolute values of each element of the matrix A, thus, $\|A\| = \|\text{abs}[A]\|$. The Discrete-Fourier Transform, or DFT, of a sequence $\{f : t \in [0, N-1]\}$ is defined here as

$$F_N(\omega) = \text{DFT}\{f\} = \frac{1}{N}\sum_{t=0}^{N-1} f(t)e^{-j\omega t}, \quad \omega \in \Omega_N = \{\frac{2\pi k}{N} : k \in [0, N-1]\}$$

where Ω_N are the normalized DFT frequencies.

Theorem 1 (Bound on Parameter Error) *Let $\{\Phi_N(\omega), Y_N(\omega), U_N(\omega) : \omega \in \Omega_N\}$ denote the DFT's of the N-point sequences $\{\phi, y, u : t \in [0, N-1]\}$, respectively, with DFT frequencies Ω_N. Let $w_{\text{BF}}(t)$ and $w_{\text{AF}}(t)$ denote the pulse response of $W_{\text{F}}\Delta_{\text{B}}$ and $W_{\text{F}}\Delta_{\text{A}}$, respectively. Suppose there are stable transfer functions W_{B} and W_{A}, and constants $M > 0$ and $\rho \in (0,1)$, such that:*

$$\begin{aligned} |\Delta_{\text{B}}(e^{j\omega})| &\leq |W_{\text{B}}(e^{j\omega})|, \ \forall|\omega| \leq \pi & (34)\\ |\Delta_{\text{A}}(e^{j\omega})| &\leq |W_{\text{A}}(e^{j\omega})|, \ \forall|\omega| \leq \pi & (35)\\ \max\{|w_{\text{BF}}(t)|, \ |w_{\text{AF}}(t)|\} &\leq M\rho^t, \ \forall t \geq 0 & (36) \end{aligned}$$

Then,

$$\|\tilde{\theta}_N^{\Delta}\| \leq \|\gamma_{\text{A}}\| + \|\gamma_{\text{B}}\| + \frac{\beta}{N} \tag{37}$$

where $\gamma_{\text{A}}, \gamma_{\text{B}} \in \mathbb{R}^p$ are given by

$$\gamma_{\text{A}} = \sum_{\omega \in \Omega_N} \text{abs}\left[W_{\text{F}}(e^{j\omega})W_{\text{A}}(e^{j\omega})Y_N(\omega)\left\langle \phi\phi^T \right\rangle_N^{-1} \Phi_N(\omega)\right] \tag{38}$$

$$\gamma_{\text{B}} = \sum_{\omega \in \Omega_N} \text{abs}\left[W_{\text{F}}(e^{j\omega})W_{\text{B}}(e^{j\omega})U_N(\omega)\left\langle \phi\phi^T \right\rangle_N^{-1} \Phi_N(\omega)\right] \tag{39}$$

and β is the constant

$$\beta = \left(\sup_{t \in [0,N-1]} \left\| \left\langle \phi\phi^T \right\rangle_N^{-1} \phi(t) \right\| \right) \left(\sup_{t \in (-\infty,-1]} \left\| \begin{pmatrix} y(t) - y_N(t) \\ u(t) - u_N(t) \end{pmatrix} \right\| \right) \left(\frac{M\rho(1-\rho^N)}{(1-\rho)^2}\right) \tag{40}$$

with $\{y_N, u_N : t \in (-\infty, \infty)\}$ the N-periodic extension of $\{y, u : t \in [0, N-1]\}$.

Proof: See Appendix.

Remarks: (1) The N-periodic extension of the N-point sequences y, u are obtained by copying y, u for all N-point intervals in $(-\infty, \infty)$.

(2) In the special case when y, u are N-periodic, $\beta = 0$, and the bound (37) on the parameter error is tight, that is,

$$\sup_{\Delta_B, \Delta_A} \|\tilde{\theta}_N^{\Delta}\| = \|\gamma_A\| + \|\gamma_B\| \tag{41}$$

Thus, for large values of N, the bound is asymptotically tight.

(3) The constant $\rho \in (0,1)$ is the magnitude of the least stable pole of either $W_F \Delta_B$ or $W_F \Delta_A$.

4.1 Computing The Bound

To compute the bound in (37) requires both measured and *a priori* data. The first two terms in the bound involve the vectors γ_B, γ_A. These depend on the measured N-point data $\{\phi, y, u : t \in [0, N-1]\}$ and on *a priori* knowledge of the model weighting functions W_B and W_A. The measured data can be used to compute the required DFT's as well as $\left\langle \phi\phi^T \right\rangle_N$.

The last term in (37) decays as $1/N$, and hence, can be neglected for a sufficiently long data record. If not neglected, then its computation depends on a number β. To compute β, observe that the first and second factors in β depend on measured data. The first uses current data, that is, data for $t \in [0, N-1]$. However, the second factor needs information from Biblical Times ($t = -\infty$) to just before the current data record ($t = 0$). Since this second factor is essentially a measure of the nearness to N-periodicity of y, u, it is likely that a reasonable upper bound can be found from analysis or experiments. The third factor depends on *a priori* knowledge of the pulse response of the nonparametric error dynamics.

4.2 Design Choices

The designer can affect the size of the bound by direct choice of the data filter W_F, and indirectly, by manipulating some exogenous *probing signals*. Both of these choices will effect γ_B, γ_A, and β.

Selection of the data filter W_F is based on the underlying assumption that the parametric model G_{θ_o} is a good approximation of G over some range of frequencies Ω_F. Hence, a natural choice for the data filter is one that has the property that

$$|W_F(e^{j\omega})| = \begin{cases} 1 + \mathcal{O}(\eta), & \omega \in \Omega_F \\ \mathcal{O}(\eta), & \text{otherwise} \end{cases} \tag{42}$$

where η is a small positive number which can be interpreted as passband or stopband "ripple". As a result we have:

Corllary 1 *If W_F satisfies (42), then for all small $\eta > 0$,*

$$\|\tilde{\theta}_N^{\Delta}\| \leq \left(\sup_{\omega \in \Omega_F} |W_A(e^{j\omega})|\right) \|\overline{\gamma}_A\| + \left(\sup_{\omega \in \Omega_F} |W_B(e^{j\omega})|\right) \|\overline{\gamma}_B\| + \mathcal{O}(\eta) + \frac{\beta}{N} \tag{43}$$

where β is from Theorem 1 and $\overline{\gamma}_A, \overline{\gamma}_B \in \mathbb{R}^p$ are given by:

$$\overline{\gamma}_A = \sum_{\omega \in \Omega_F} \text{abs}\left[Y_N(\omega) \left\langle \phi\phi^T \right\rangle_N^{-1} \Phi_N(\omega)\right] \tag{44}$$

$$\overline{\gamma}_B = \sum_{\omega \in \Omega_F} \text{abs}\left[U_N(\omega) \left\langle \phi\phi^T \right\rangle_N^{-1} \Phi_N(\omega)\right] \tag{45}$$

For large N and small η, the bound in (43) is proportional to the peak magnitude in the frequency domain of the nonparametric dynamics in the passband Ω_F. This is unlike robust control where a bound is needed over *all* frequencies. Thus, the data filter can be used to offset the impact of the poorly modeled nonparametric dynamics. A typical example is

$$\Omega_F = \{\omega : |\omega| \in [\omega_a, \omega_b]\} \tag{46}$$

The bound in (43) is also proportional to a sum of vector norms ($\|\overline{\gamma}_A\| + \|\overline{\gamma}_B\|$) which can be related to the *condition number* of the matrix $\left\langle \phi\phi^T \right\rangle_N$. Recall that

$$\text{cond}\left\{\left\langle \phi\phi^T \right\rangle_N\right\} \stackrel{\text{def}}{=} \|\left\langle \phi\phi^T \right\rangle_N^{-1}\| \, \|\left\langle \phi\phi^T \right\rangle_N\| \tag{47}$$

To see the relation, observe that

$$\begin{aligned} \|\overline{\gamma}_A\| &\leq \|\left\langle \phi\phi^T \right\rangle_N^{-1}\| \, \|\sum_{\omega \in \Omega_N} \text{abs}\left[Y_N(\omega)\Phi_N(\omega)\right]\| \\ &= \text{cond}\left\{\left\langle \phi\phi^T \right\rangle_N\right\} \left(\frac{\|\sum_{\omega \in \Omega_N} \text{abs}\left[Y_N(\omega)\Phi_N(\omega)\right]\|}{\|\langle \phi\phi^T \rangle_N\|}\right) \end{aligned}$$

Clearly the second factor above is near unity because ϕ consists of filtered versions of y, u. Thus, data which results in a large condition number is just as bad as large nonparametric model error in the passband.

Observe that the smallest value of the condition number is one, and minimization of it is a long standing difficult problem, especially in closed-loop where $\left\langle \phi\phi^T \right\rangle_N$ depends in a complicated manner on the true plant dynamics and all the exogenous inputs, see, *e.g.*, Mareels *et al.* (1986).

5 Application to a Special Case

To provide a clearer interpretation of Theorem 1 and Corollary 1, we specialize the results to the model structure of (7), that is, where

$$G = \frac{\sum_{i=1}^{m} b_{o,i} q^{-i}}{1 + \sum_{i=1}^{m} a_{o,i} q^{-i}} L \tag{48}$$

$$L = \left[1 + \widetilde{L}\right] \widehat{L} \tag{49}$$

and where $\widetilde{L}$ represents uncertain stable nonparametric dynamics known to be bounded by

$$|\widetilde{L}(e^{j\omega})| \leq |W_L(e^{j\omega})|, \ \forall |\omega| \leq \pi \tag{50}$$

We then have:

$$B_{\theta_o} = \sum_{i=1}^{m} b_{o,i} q^{-i} \widehat{L} \tag{51}$$

$$A_{\theta_o} = 1 + \sum_{i=1}^{n} a_{o,i} q^{-i} \tag{52}$$

$$\Delta_{\mathrm{B}} = \widetilde{L} \sum_{i=1}^{m} b_{o,i} q^{-i} \widehat{L} \tag{53}$$

$$\Delta_{\mathrm{A}} = 0 \tag{54}$$

The least-squares parameter error due to the unmodeled dynamics and distrbance is in this case:

$$\widetilde{\theta}_N^{\Delta} = \left\langle \phi\phi^T \right\rangle_N^{-1} \langle \phi \varepsilon_\Delta \rangle_N \tag{55}$$

$$\widetilde{\theta}_N^{d} = \left\langle \phi\phi^T \right\rangle_N \langle \phi \varepsilon_d \rangle_N \tag{56}$$

$$\varepsilon_\Delta = \widetilde{L} \sum_{i=1}^{m} b_{o,i} q^{-i} \widehat{L} u \tag{57}$$

$$\varepsilon_d = W_{\mathrm{F}} A_{\theta_o} d \tag{58}$$

$$\phi = W_{\mathrm{F}} \widehat{L} \left(-q^{-1} y \cdots - q^{-n} y \ q^{-1} u \cdots q^{-m} u\right)^T \tag{59}$$

The above definitions lead to the more concise expression:

$$\begin{aligned} \varepsilon_\Delta &= \widetilde{L} \phi^T \begin{pmatrix} 0_{n\times 1} \\ b_o \end{pmatrix} \\ b_o^T &= (b_{o,1} \cdots b_{o,m}) \end{aligned} \tag{60}$$

Thus,

$$\widetilde{\theta}_N^{\Delta} = \left\langle \phi\phi^T \right\rangle_N^{-1} \left\langle \phi (\widetilde{L}\phi)^T \right\rangle_N \begin{pmatrix} 0_{n\times 1} \\ b_o \end{pmatrix} \tag{61}$$

As a result, following Theorem 1 and Corollary 1, we can state:

Theorem 2

$$\|\tilde{\theta}_N^{\Delta}\| \leq \|\Gamma\| \, \|b_o\| + \mathcal{O}(1/N) \tag{62}$$

where $\Gamma \in \mathbb{R}^{p \times m}$ *is:*

$$\Gamma = \sum_{\omega \in \Omega_N} \text{abs}\left[W_L(e^{j\omega}) \left\langle \phi\phi^T \right\rangle_N^{-1} \Phi_N(\omega)\Phi_N(\omega)^T \begin{pmatrix} 0_{n \times m} \\ I_m \end{pmatrix} \right] \tag{63}$$

In addition, if W_F *satisfies (42), then*

$$\|\tilde{\theta}_N^{\Delta}\| \leq \left(\sup_{\omega \in \Omega_F} |W_L(e^{j\omega})| \right) \|\overline{\Gamma}\| \, \|b_o\| + \mathcal{O}(\eta) + \mathcal{O}(1/N) \tag{64}$$

where

$$\overline{\Gamma} = \sum_{\omega \in \Omega_F} \text{abs}\left[\left\langle \phi\phi^T \right\rangle_N^{-1} \Phi_N(\omega)\Phi_N(\omega)^T \begin{pmatrix} 0_{n \times m} \\ I_m \end{pmatrix} \right] \tag{65}$$

An example of this model structure is presented next.

5.1 Example: Rotating Flexible Servo

The example system is a rotating flexible servo described in Åströmand Wittenmark(1988). It consists of a load connected by a flexible shaft to a motor/tachometer. The servo design objective is to control the motor velocity despite changing loads and uncertain shaft flexibilty. It is assumed that the flexible shaft is adequately modeled by its first flexible mode whose frequency is well below the bandwidth of the motor/tach electronics. Thus, a good continuous-time transfer function model from motor control voltage to motor/tach velocity output is[1]

$$G(s) = \frac{K_E(J_L s^2 + D_S s + K_S)}{s[J_L J_M s^2 + (J_L + J_M)D_S s + (J_L + J_M)K_S]} \tag{66}$$

J_L is the load inertia, J_M is the motor/tach inertia, K_S is the shaft stiffness, D_S is the shaft damping, and K_E is the product of the motor and tach electronic gains.

5.1.1 Parameter Set Estimation

The parametric model used for parameter set estimation is obtained by neglecting the shaft flexibility, thus,

$$G_{\theta_o}(s) = \frac{\theta_o}{s} \tag{67}$$

[1]Beware that we use the same nomenclature to denote continuous-time and discrete-time transfer functions, *e.g.*, $G(s)$ and $G(q)$. These are *not* to be read as the same function of a different variable s and q. In this case, $G(q) \stackrel{\text{def}}{=} \mathcal{ZOH}\{G(s)\}$ where $\mathcal{ZOH}\{\cdot\}$ is the usual zero-order-hold transformation.

with true parameter

$$\theta_o = \frac{K_E}{J_L + J_M} \tag{68}$$

The effect of the changing load appears as a change in the parameter θ_o. The discrete-time parametric transfer function is then

$$G_{\theta_o}(q) = \mathcal{ZOH}\left\{\frac{\theta_o}{s}\right\} = \frac{\theta_o h q^{-1}}{1-q^{-1}} \tag{69}$$

where h is the sampling interval and $\mathcal{ZOH}\{\cdot\}$ is the usual zero-order-hold transformation. By neglecting the flexible modes, we have implicitly made the selection

$$\hat{L} = 1 \tag{70}$$

Hence,

$$\begin{aligned} B_{\theta_o} &= \theta_o q^{-1} && (71)\\ A_{\theta_o} &= \frac{1}{h}(1-q^{-1}) && (72)\\ \Delta_{\mathrm{A}} &= 0 && (73)\\ \Delta_{\mathrm{B}} &= \theta_o q^{-1}\tilde{L} && (74) \end{aligned}$$

where the nonparametric transfer function is

$$\begin{aligned} \tilde{L} &= \frac{\mathcal{ZOH}\{G(s) - G_{\theta_o}(s)\}}{\mathcal{ZOH}\{G_{\theta_o}(s)\}} && (75)\\ &= \frac{1-q^{-1}}{hq^{-1}}\,\mathcal{ZOH}\left\{\frac{\left(\frac{J_L^2}{J_L+J_M}\right)s}{\left(\frac{J_L J_M}{J_L+J_M}\right)s^2 + D_S s + K_S}\right\} && (76) \end{aligned}$$

Since the denominator delay will be cancelled by a numerator delay, $\tilde{L}$ is stable [see any table of $\mathcal{ZOH}\{\cdot\}$-transforms]. Also, $\tilde{L}(1) = 0$, thereby reinforcing the assumption that G_{θ_o} is a good low frequency model.

The parametric equation error (16) is now

$$\varepsilon_\theta = z - \theta\phi$$

with

$$\begin{aligned} z &= W_{\mathrm{F}}\frac{1}{h}(1-q^{-1})y\\ \phi &= W_{\mathrm{F}} q^{-1} u \end{aligned}$$

From Lemma 1 we have,

$$\begin{aligned} z &= \theta_o\phi + \varepsilon_\Delta + \varepsilon_d && (77)\\ \varepsilon_\Delta &= \theta_o\tilde{L}\phi && (78)\\ \varepsilon_d &= W_{\mathrm{F}}\frac{1}{h}(1-q^{-1})d && (79) \end{aligned}$$

The parameter estimate is then

$$\widehat{\theta} = \frac{\langle \phi z \rangle_N}{\langle \phi^2 \rangle_N} \tag{80}$$

Effect of Nonparametric Dynamics

The parameter error due to the nonparametric dynamics is

$$\tilde{\theta}_N^{\Delta} = \theta_o \frac{\left\langle \phi(\tilde{L}\phi) \right\rangle_N}{\langle \phi^2 \rangle_N} \tag{81}$$

From Theorem 2 we have the bound:

$$\left| \frac{\tilde{\theta}_N^{\Delta}}{\theta_o} \right| \leq \delta_1 + \mathcal{O}(1/N) \tag{82}$$

where

$$\begin{aligned} \delta_1 &= \frac{\sum_{\omega \in \Omega_N} |\tilde{L}(e^{j\omega})| \cdot |\Phi_N(\omega)|^2}{\sum_{\omega \in \Omega_N} |\Phi_N(\omega)|^2} \\ \Phi_N(\omega) &= \mathrm{DFT}\{\phi\} = \mathrm{DFT}\left\{ W_{\mathrm{F}} q^{-1} u \right\}, \ \ \omega \in \Omega_N \end{aligned}$$

If, in addition, W_{F} satisfies (42), then

$$\delta \leq \delta_2 + \mathcal{O}(\eta) \tag{83}$$

where

$$\delta_2 = \sup_{\omega \in \Omega_{\mathrm{F}}} |\tilde{L}(e^{j\omega})|$$

Simulation Results

Simulations were performed using the following values from Åström and Wittenmark(1988):

$$\begin{aligned} J_L &\in [.0002, .002] \\ J_M &= .002 \\ K_E &= .5 \\ K_S &= 100. \\ D_S &= .0001 \end{aligned} \tag{84}$$

The frequency and damping ratio of the flexible mode, as well as the true parameter θ_o vary with $J_L \in [.0002, .002]$ as follows:

$$\begin{aligned} \text{frequency} &= \frac{1}{2\pi}\sqrt{\frac{(J_L + J_M)K_S}{J_L J_M}} \in [118, 50] \text{ (hz)} \\ \text{damping} &= \frac{D_S}{2}\sqrt{\frac{J_L + J_M}{K_S J_L J_M}} \in [1.582 \times 10^{-4}, 1.581 \times 10^{-4}] \\ \theta_o &\in [125, 227.27] \end{aligned}$$

Suppose the sampling interval is $h = .002$ sec, thus, the sampling frequency is 500 hz. Figure 2 shows a plot of $|\tilde{L}(e^{j\omega})|$, $\omega = 2\pi hf$, vs. f in hz for varying $J_L \in [.0002, .002]$.

Suppose that the true but unknown load is $J_L = .002$. Thus, the unknown parameter is $\theta_o = 125$ which is to be estimated. Since the plant is unstable we use the stabilizing, but low performance, control

$$u = 1.5(r - y) \tag{85}$$

where $r(t)$ is a user applied reference command.

Single Tone Reference In this experiment the reference is the single tone

$$r(t) = \sin(\omega_r t),\ \omega_r = 2\pi h f_r \tag{86}$$

and the data filter is

$$W_F = 1 \tag{87}$$

The parameter error bound is then

$$\delta = \frac{\sum_{\omega\in\Omega_N} |\tilde{L}(e^{j\omega})| \cdot |\Phi_N(\omega)|^2}{\sum_{\omega\in\Omega_N} |\Phi_N(\omega)|^2} = |\tilde{L}(e^{j\omega_r})| + \mathcal{O}(1/N) \tag{88}$$

Simulations were performed with the system initially at rest with input tones at

$$f_r = [5, 10, 20, 40] \text{ hz}$$

At each tone a BLS parameter estimate was obtained for data records of length

$$N = [64, 128, 256, 512, 1024]$$

The results are shown in Figures 3-5. Figures 3 and 4 show, respectively, the true normalized parameter error $|\tilde{\theta}_N^{\Delta}/\theta_o|$ and the parameter error estimate δ from (88) superimposed (in $\times$'s) on the graph of $|\tilde{L}(e^{j\omega})|$ at each reference frequency ω_r for varying data record length N. Observe that as N increases, both the error and the bound approach the asymptotic upper bound $|\tilde{L}(e^{j\omega_r})|$ corresponding to the input tone. This is predicted by the theory. Figure 5 displays the same data as a function of N, *i.e.*, $|\tilde{\theta}_N^{\Delta}/\theta_o|$, the bound δ, and the limiting value $|\tilde{L}(e^{j\omega_r})|$ which does not depend on N. All the values are converging to this limit as N increases. A summary of the plotted data is also contained in Table 1.

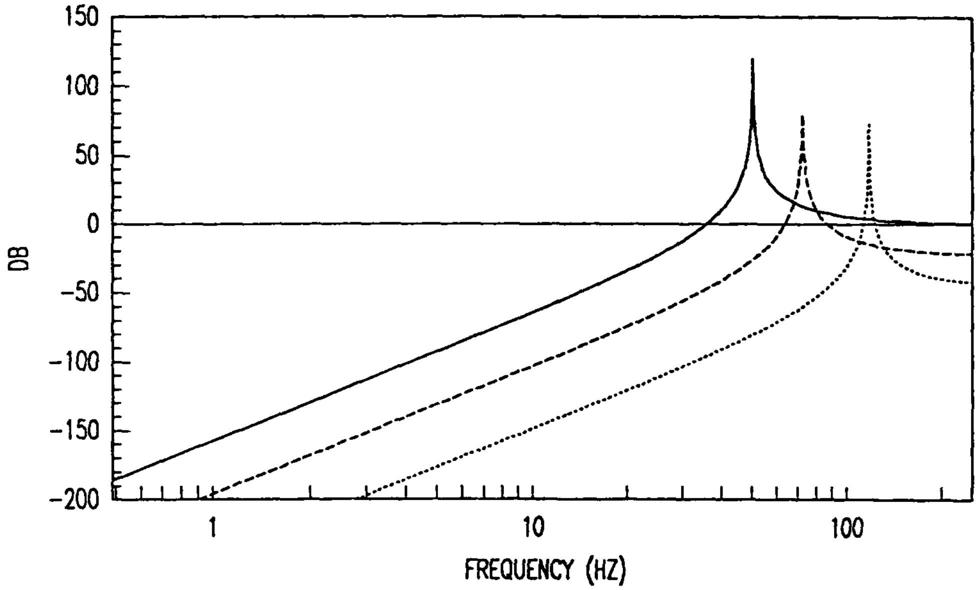

Figure 2: $|\tilde{L}(e^{j\omega})|$ vs. ω for varying $J_L = .002$ (dark line), $J_L = .00063$ (dashed line), and $J_L = .0002$ (dotted line).

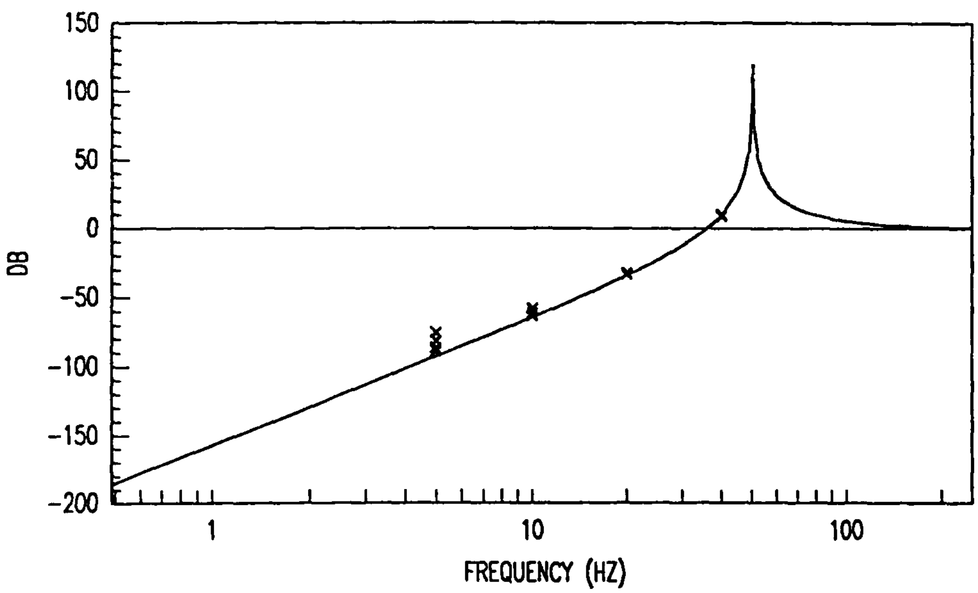

Figure 3: The $\times$'s are the true normalized parameter error $|\tilde{\theta}_N^{\Delta}/\theta_o|$ superimposed on the graph of $|\tilde{L}(e^{j\omega})|$ vs. ω (dark line) at each reference frequency ω_r for varying data record length N.

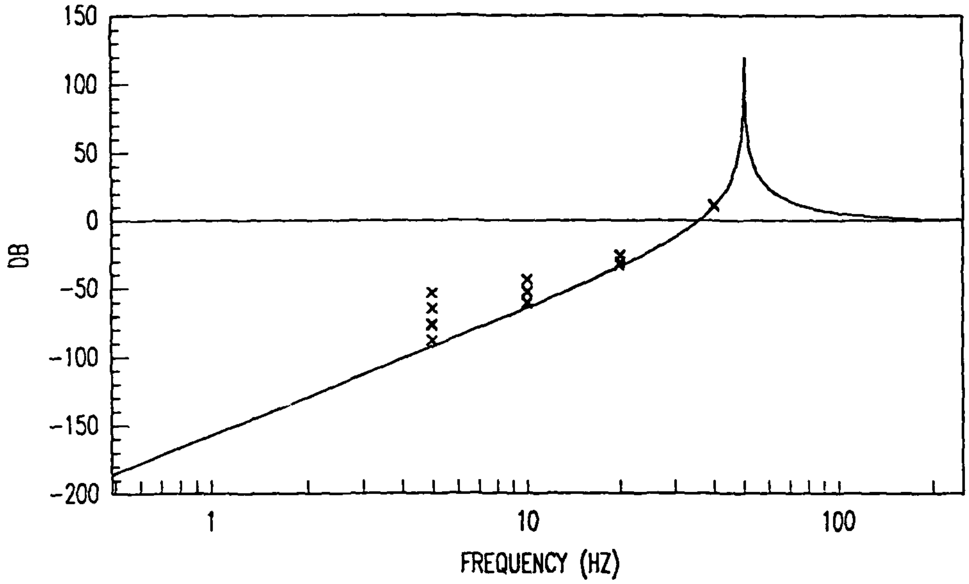

Figure 4: The $\times$'s are the values of the parameter error estimate δ from (88) superimposed on the graph of $|\tilde{L}(e^{j\omega})|$ vs. ω (dark line) at each reference frequency ω_r for varying data record length N.

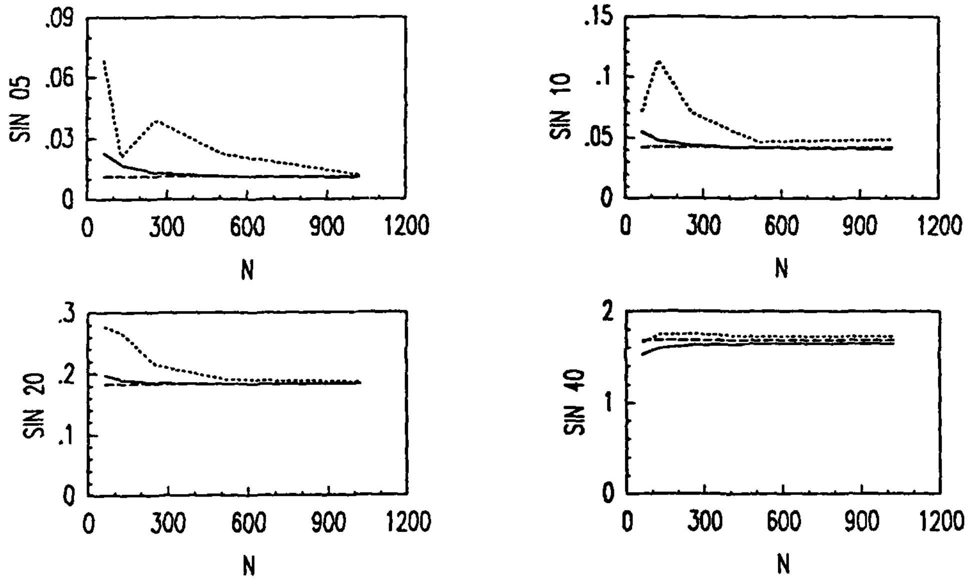

Figure 5: Plotted vs. data length N are: normalized parameter error $|\tilde{\theta}_N^{\Delta}/\theta_o|$ (dark line), parameter error estimate δ (dotted line), and $|\tilde{L}(e^{j\omega_r})|$ (dashed line).

f_r	N	$\lvert\widetilde{\theta}_N^{\triangle}/\theta_o\rvert$	δ	$\lvert\widetilde{L}(e^{j\omega_r})\rvert$
5	64	.0228	.0684	.0112
	128	.0167	.0208	
	256	.0131	.0389	
	512	.0113	.0222	
	1024	.0105	.0119	
10	64	.0544	.0714	.0420
	128	.0470	.1132	
	256	.0433	.0699	
	512	.0415	.0465	
	1024	.0406	.0483	
20	64	.1976	.2771	.1825
	128	.1888	.2661	
	256	.1849	.2131	
	512	.1831	.1902	
	1024	.1822	.1857	
40	64	1.5267	1.6631	1.6848
	128	1.5971	1.7490	
	256	1.6291	1.7562	
	512	1.6438	1.7243	
	1024	1.6510	1.7240	

Table 1: Summary of simulation results with single tone reference

Square-Wave Reference We repeat the experiment with a unit amplitude square-wave refrence whose base period is $256h = .512$ sec (1.953 hz). The results are shown in Figures 6-9. Figure 6 shows the time histories of the plant input and output for 1024 data points, *i.e.,* $1023h = 2.046$ sec. The data filter is an 8th order Butterworth low-pass filter with cutoff frequency ω_F. The experiments correspond to the following values of cutoff frequency:

$$\omega_F = 2\pi h f_F, \; f_F = [5, 10, 20, 40] \text{ hz}$$

The filter then behaves like (42) with the passband

$$\Omega_F = \{\omega \in \Omega_N : |\omega| \leq \omega_F\} \tag{89}$$

Figures 7 and 8 show, respectively, the true normalized parameter error $|\tilde{\theta}_N^\Delta/\theta_o|$ and the bound δ_1 from (82) (the $\times$'s) superimposed on the graph of $|\tilde{L}(e^{j\omega})|$ at each filter cutoff frequency ω_F for varying data record length N.

Observe that as N increases, both the true error and the estimate approach the asymptotic limit $|\tilde{L}(e^{j\omega_F})|$ corresponding to the filter cutoff frequency. This is predicted by the theory. Figure 9 displays the same data as a function of N, *i.e.,* $|\tilde{\theta}_N^\Delta/\theta_o|$, the bound δ, and the constant limiting value $|\tilde{L}(e^{j\omega_F})|$. In this case all the values are below or near the limit except for one instance at $\omega_F = 5$hz and a very short data record of $N = 64$. A summary of all the square-wave data is contained in Table 2.

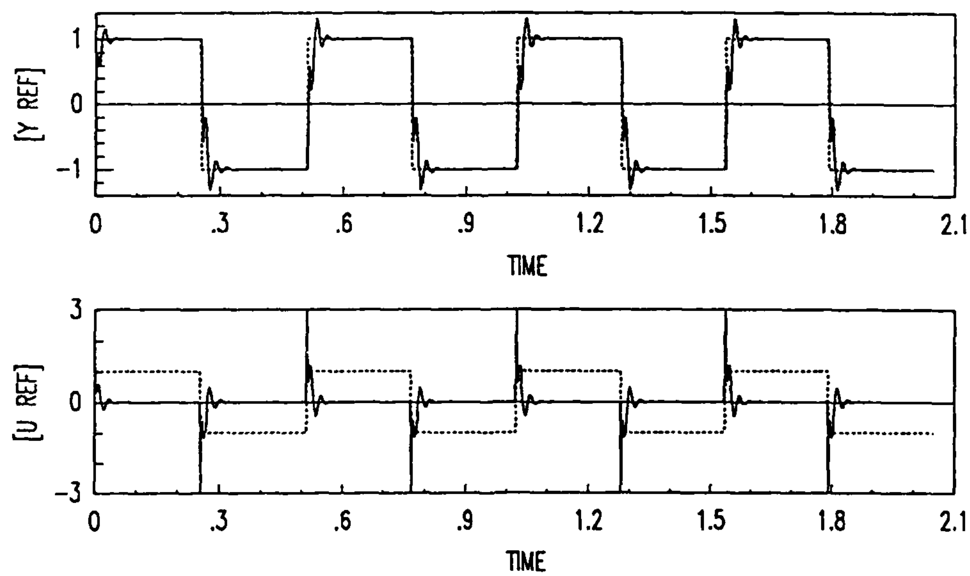

Figure 6: Time histories of (y, r) (upper) and (u, r) (lower).

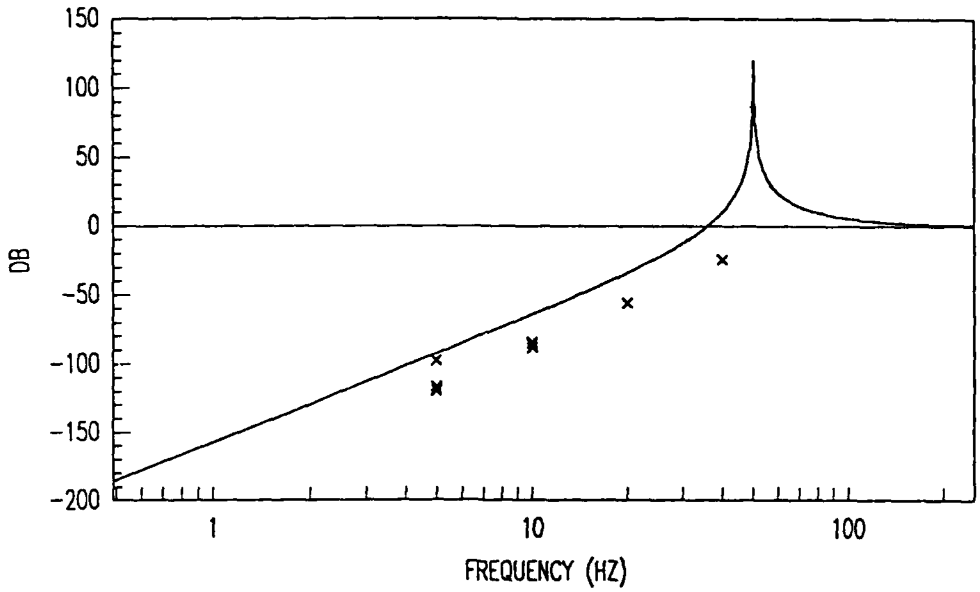

Figure 7: The $\times$'s are the true normalized parameter error $|\tilde{\theta}_N^{\triangle}/\theta_o|$ superimposed on the graph of $|\tilde{L}(e^{j\omega})|$ vs. ω (dark line) at each filter frequency ω_F for varying data record length N.

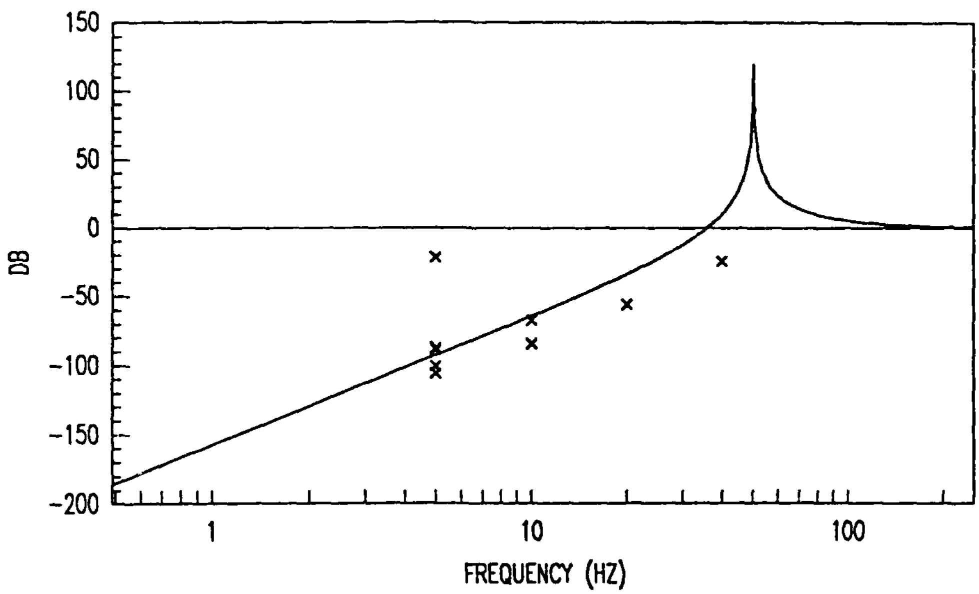

Figure 8: The $\times$'s are the values of the parameter error estimate δ from (88) superimposed on the graph of $|\tilde{L}(e^{j\omega})|$ vs. ω (dark line) at each filter frequency ω_F for varying data record length N.

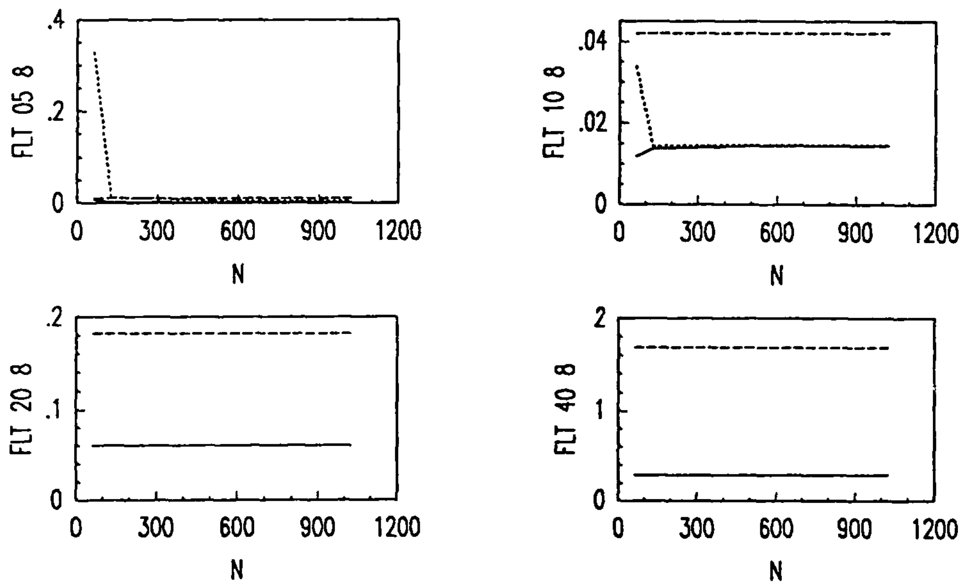

Figure 9: Plotted vs. data length N are: normalized parameter error $|\tilde{\theta}_N^{\Delta}/\theta_o|$ (dark line), parameter error estimate δ (dotted line), and $|\tilde{L}(e^{j\omega_F})|$ (dashed line).

f_F	N	$\lvert\tilde{\theta}_N^{\Delta}/\theta_o\rvert$	δ	$\lvert\tilde{L}(e^{j\omega_F})\rvert$
5	64	.0074	.3289	.0112
	128	.0029	.0126	
	256	.0026	.0114	
	512	.0024	.0064	
	1024	.0023	.0050	
10	64	.0119	.0338	.0420
	128	.0137	.0144	
	256	.0140	.0145	
	512	.0143	.0146	
	1024	.0144	.0146	
20	64	.0597	.0604	.1825
	128	.0604	.0607	
	256	.0603	.0606	
	512	.0603	.0605	
	1024	.0603	.0605	
40	64	.2882	.2906	1.6848
	128	.2882	.2906	
	256	.2882	.2906	
	512	.2882	.2906	
	1024	.2882	.2906	

Table 2: Summary of simulation results with square wave reference

5.1.2 Set-Point Regulator Design

The problem is to design a set point regulator for the true plant

$$G = G_{\theta_o}(1+\tilde{L}),\; G_{\theta_o} = \frac{\theta_o h q^{-1}}{1-q^{-1}}$$

where θ_o and $\tilde{L}$ are unknown elements of the sets:

$$\begin{aligned} \theta_o &\in [125, 227.27] \\ |\tilde{L}(e^{j\omega})| &\leq |W_L(e^{j\omega})|,\; \forall|\omega| \leq \pi \end{aligned}$$

where W_L is known. We will take

$$W_L = \tilde{L}\,|_{J_L=.002}$$

which is the worst case as seen in Figure 2. The regulator is based on the parameter set estimator of θ_o, where

$$|\hat{\theta} - \theta_o| \leq \delta|\theta_o|$$

with $\hat{\theta}$ from (80) and δ given by either (82) or (83). The control sequence is obtained via the feedback system

$$u = K_{\hat{\theta}}(r-y)$$

where $K_{\hat{\theta}}$, the regulator transfer function, is selected to satisfy

$$\frac{K_{\hat{\theta}}G_{\hat{\theta}}}{1+K_{\hat{\theta}}G_{\hat{\theta}}} = T_*$$

with T_* the desired closed-loop trnasfer function from r to y. Observe that T_* does not depend on the parameter estimate $\hat{\theta}$. Applying this control to the actual plant results in the closed-loop transfer function

$$T = T_* \frac{(1+\tilde{G})(1+\tilde{L})}{1+T_*\widetilde{M}}$$

where

$$\begin{aligned} \tilde{G} &= \frac{G_{\theta_o}}{G_{\hat{\theta}}} - 1 \\ \widetilde{M} &= (1+\tilde{G})(1+\tilde{L}) - 1 \end{aligned}$$

The transfer function $\tilde{G}$ reflects the parameter estimation error, and in this example reduces to the constant

$$\tilde{G} = \frac{\theta_o}{\hat{\theta}} - 1$$

Thus,

$$\widetilde{M} = (1+\tilde{L})\frac{\theta_o}{\hat{\theta}} - 1 = \frac{\tilde{L} - \tilde{\theta}_N^{\Delta}/\theta_o}{1+\tilde{\theta}_N^{\Delta}/\theta_o}$$

where $\widetilde{\theta}_N^{\Delta}$ is the parameter error induced by the unmodeled dynamics $\widetilde{L}$. Since $\widetilde{G}, \widetilde{L}$, and T_* are all stable, it follows that a sufficient condition for T to be stable is

$$|T_*(e^{j\omega})\widetilde{M}(e^{j\omega})| < 1, \ \forall|\omega| \leq \pi \tag{90}$$

To provide a numerical example, consider the experiment with a square-wave reference and a low pass data filter with $f_F = 20$ hz. With $N = 1024$, we have $|\widetilde{\theta}_N^{\Delta}/\theta_o| = .0603$, $\delta_1 = .0605$, and $\delta_2 = .1825$ [see Table 2]. Figure 10 shows $|\widetilde{M}(e^{j\omega})|$ vs. ω for normalized parameter errors of $\widetilde{\theta}_N^{\Delta}/\theta_o = \pm.0603, \pm.1825$. The trouble spot is clearly at 50 hz where the error is about 125 db. Suppose that T_* is chosen as a 4th order 10hz low pass filter. Figure ?? shows that the closed-loop system is stable since the above inequality is satisfied.

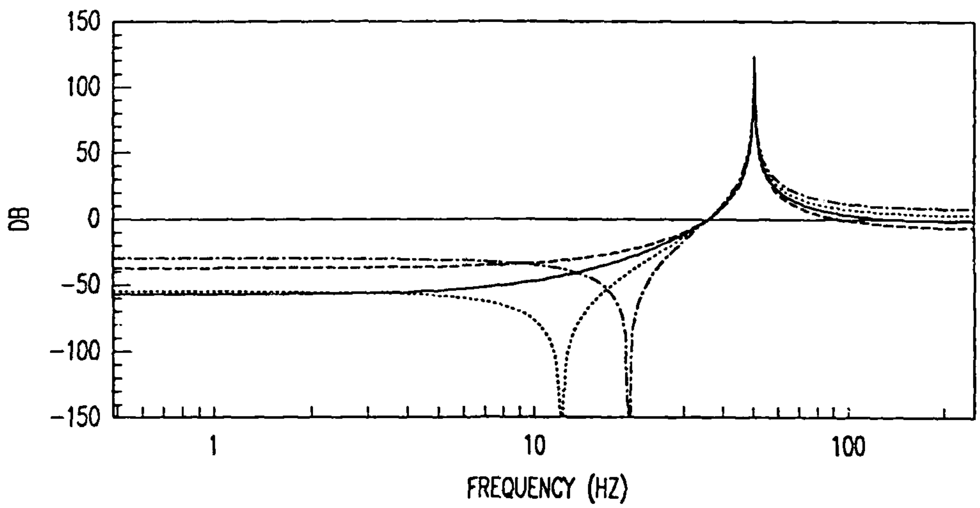

Figure 10: $|\widetilde{M}(e^{j\omega})|$ vs. ω for normalized parameter errors of $\widetilde{\theta}_N^{\Delta}/\theta_o = +.0603$ (dark line), -.0603 (dotted line), +.1825 (dashed line), -.1825 (dash-dotted line).

Further restrictions are required to satisfy performance and not just stability. A convenient expression which reflects closed-loop performance is the relative error between the true closed-loop T and the desired T_*, that is:

$$\widetilde{T} \stackrel{\text{def}}{=} \frac{T}{T_*} - 1 = \frac{(1 - T_*)\widetilde{M}}{1 + T_*\widetilde{M}}$$

Suppose that the closed-loop performance tolerance is

$$|\widetilde{T}(e^{j\omega})| \leq \epsilon(\omega), \ \forall|\omega| \leq \pi$$

where $\epsilon(\omega)$ is a specified function of frequency. A typical example is

$$\epsilon(\omega) = \begin{cases} \epsilon_{\text{low}}, & |\omega| \leq \omega_c \\ \epsilon_{\text{high}}, & |\omega| > \omega_c \end{cases}$$

The frequency ω_c represents the closed-loop performance bandwidth, *e.g.,* $T_*(e^{j\omega}) \approx 1, |\omega| \leq \omega_c$. A sufficient condition to insure the performance constraint is that

$$\sup_{|\hat{\theta}-\theta_o| \leq \delta |\theta_o|} |\widetilde{M}(e^{j\omega})| \leq \ell(\omega) \stackrel{\text{def}}{=} \frac{\epsilon(\omega)}{|1 - T_*(e^{j\omega})| + \epsilon(\omega)|T_*(e^{j\omega})|}$$

which also insures that $|T_*(e^{j\omega})\widetilde{M}(e^{j\omega})| < 1$, thus, T is stable. Using in place of the left hand side above an upper bound results in

$$\frac{|\tilde{L}(e^{j\omega})| + \delta}{1 - \delta} \leq \ell(\omega)$$

Finally, an upper bound on parameter error, sufficient to insure performance as defined above, is

$$\delta \leq \delta_m \stackrel{\text{def}}{=} \min_{|\omega| \leq \pi} \frac{\ell(\omega) - |\tilde{L}(e^{j\omega})|}{1 + \ell(\omega)}$$

Observe that for consisitancy we need

$$|\tilde{L}(e^{j\omega})| < \ell(\omega)$$

which also enforces the requirement that $\delta < 1$. To see how to use these relations for design, let $\epsilon_{\text{low}} = .1$, $\epsilon_{\text{high}} = 10,000$. Figure 12 shows a plot of $|\tilde{L}(e^{j\omega})|$ and $\ell(\omega)$ vs. ω. Thus, performance will be satisfied if

$$|\tilde{\theta}_N^{\Delta}/\theta_o| \leq \delta_m = .0164$$

The minimum occurs at 10 hz. Since this value is smaller than any of the paremeter error bounds (including the actual error), we are led to lowering the data filter bandwidth. Reducing from 20 hz to 10 hz results in $|\tilde{\theta}_N^{\Delta}/\theta_o| = .0144$, $\delta_1 = .0146$, and $\delta_2 = .0420$ [see Table 2]. The measured bound δ_1 is smaller than δ_m, and hence, the closed- loop performance will be satisfactory under these conditions. Now, the actual bound at any one time could be different, but the robust control design procedure can have an option of relaxing the performance criterion.

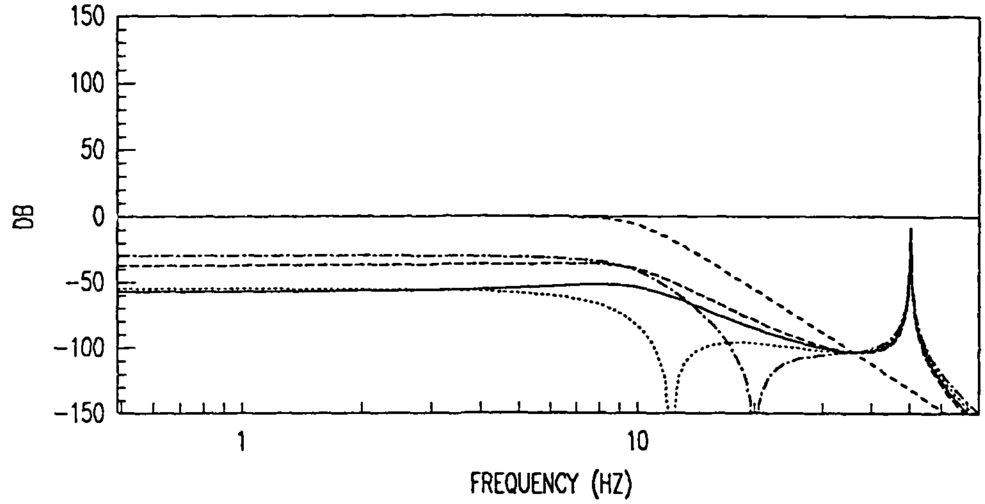

Figure 11: Plotted against ω are $|T_*(e^{j\omega})\widetilde{M}(e^{j\omega})|$ for the parameter error values in Figure 10, and $T_*(e^{j\omega})|$.

lt2

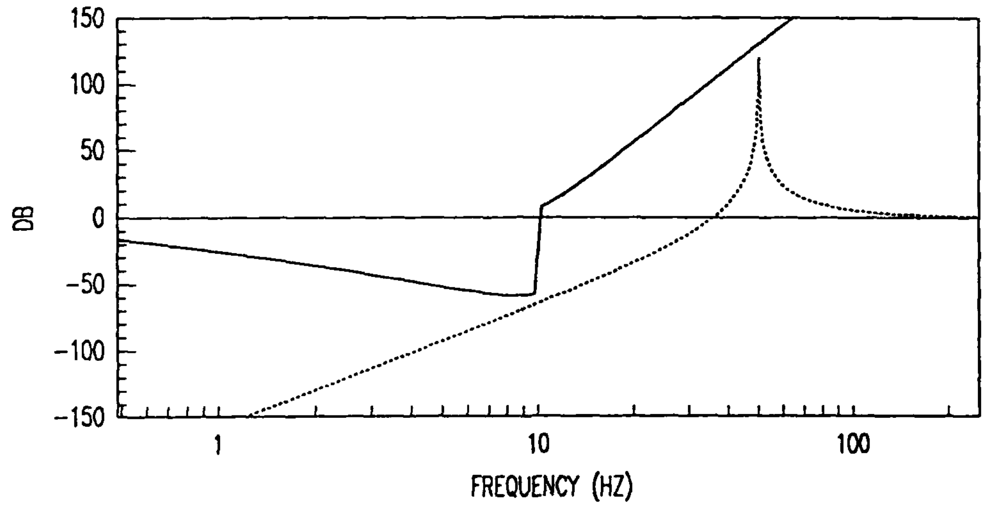

Figure 12: $|\widetilde{L}(e^{j\omega})|$ (dashed line) and $\ell(\omega)$ (dark line) vs. ω.

6 Bound on Parameter Error from Disturbance

To bound $\|\tilde{\theta}_N^d\|$ will require some *a priori* data about the disturbance sequence ε_d. Consider the special case of Section 5, with (58):

$$\varepsilon_d = W_F A_{\theta_o} d$$

The following result is obtained if ε_d has a known RMS bound.

Theorem 3 (RMS Bounded Disturbance) *Suppose that*

$$\left\langle \varepsilon_d^2 \right\rangle_N^{1/2} \leq \sigma_d \tag{91}$$

Then:

$$\|\tilde{\theta}_N^d\| \leq \sigma_d \left\langle \| \left\langle \phi\phi^T \right\rangle_N^{-1} \phi\|^2 \right\rangle_N^{1/2} \tag{92}$$

Proof: Direct application of Cauchy-Schwartz and triangle inequalities.

Use of this bound requires prior knowledge of σ_d. Rather than producing this value analytically, since ε_d depends only on the true plant and disturbance, it is easier to run some simulations or perform some experiments.

The effect of the disturbance on the parameter error can be neglected whenever $\|\tilde{\theta}_N^d\| << \|\tilde{\theta}_N^\Delta\|$. From Lemma 1, this occurs if $\langle \varepsilon_d^2 \rangle_N << \langle \varepsilon_\Delta^2 \rangle_N$. Using the special structure in (60) gives:

$$\begin{aligned}
\left\langle \varepsilon_\Delta^2 \right\rangle_N^{1/2} &\leq \left\langle [\tilde{L}\phi^T \begin{pmatrix} 0_{n\times 1} \\ b_o \end{pmatrix}]^2 \right\rangle_N^{1/2} \\
&\leq \|b_o\| \left\langle \|\tilde{L}\phi\|^2 \right\rangle_N^{1/2} \\
&\leq \|b_o\| \left\langle \|W_L\phi\|^2 \right\rangle_N^{1/2} + \mathcal{O}\left(1/\sqrt{N}\right)
\end{aligned}$$

where the last line follows from Lemma-DFT in the Appendix. Now if W_L satisfies (42), then for large N and small η,

$$\left\langle \varepsilon_\Delta^2 \right\rangle_N \leq \|b_o\| \left(\sup_{\omega\in\Omega_F} |W_L(e^{j\omega})| \right) \left\langle \|\phi\|^2 \right\rangle_N^{1/2} + \mathcal{O}\left(1/\sqrt{N}\right) + \mathcal{O}(\eta) \tag{93}$$

Thus, the unmodeled dynamics will be the dominant contributor to parameter error if

$$\sqrt{\frac{\langle \varepsilon_d^2 \rangle_N}{\langle \|\phi\|^2 \rangle_N}} << \|b_o\| \left(\sup_{\omega\in\Omega_F} |W_L(e^{j\omega})| \right) \tag{94}$$

An interpretation is that the "noise to signal ratio" (the left hand side above) should be sufficiently small as determined by the "model error" in the passband (the right hand side above).

If the noise is known to be stochastic, then the above results tend to be conservative, since information is neglected. For example, suppose it has been determined (by simulations and experiments) that ε_d is a zero-mean sequence with auto-spectrum

$$S_{\varepsilon_d}(\omega) = \sigma_e^2 |H(e^{j\omega})|^2 \tag{95}$$

where H is a stable and stably invertible spectral factor of the standard form $H(q) = 1 + \sum_{k=1}^{\infty} h(k)q^{-k}$ [see, *e.g.*, Ljung(1987)]. Hence, ε_d is the output of the filter H whose input is a zero-mean white noise sequence e with variance σ_e^2. Although H and σ_e are unknown, suppose they are known to be bounded as follows:

$$\begin{aligned} \sigma_e &\leq \widehat{\sigma}_e \\ |W_{\mathrm{F}}(e^{j\omega})H(e^{j\omega})| &\leq |\widehat{H}(e^{j\omega})|, \ \forall|\omega| \leq \pi \end{aligned}$$

With this information available, it is more appropriate to use the equation error

$$\varepsilon_\theta = \widehat{H}^{-1} W_{\mathrm{F}}(A_\theta y - B_\theta u) \tag{96}$$

This leads to the parameter error

$$\tilde{\theta}_N^d = \left\langle \phi\phi^T \right\rangle_N^{-1} \langle \phi\varepsilon_d' \rangle_N \tag{97}$$

$$\varepsilon_d' = \widehat{H}^{-1} W_{\mathrm{F}} H e \tag{98}$$

Hence, in the passband of the data filter ε_d' is effectively zero-mean white noise with maximum variance $\widehat{\sigma}_e^2$. Taking expectations with respect to the statistics of e, it follows that

$$\mathcal{E}\{\tilde{\theta}_N^d(\tilde{\theta}_N^d)^T\} \approx \sigma_e^2 \left\langle \phi\phi^T \right\rangle_N^{-1} \tag{99}$$

Using the above *a priori* statistical data motivates the bound:

$$\|\tilde{\theta}_N^d\| \leq \widehat{\sigma}_e \| \left\langle \phi\phi^T \right\rangle_N^{-1} \|^{1/2} \tag{100}$$

This is clearly a tighter bound than that obtained in Theorem 3, which is using much less statistical data about the disturbance sequence.

7 Discussion

The parameter set estimator developed here can be used either off-line or on- line. Most of the discussion up to now focuses on one N-point fixed observation interval $t \in [0, N-1]$. In practice, the scheme would be implemented more or less as follows. First solve the least-squares parameter estimation problem for a moving N-point observation interval, say, $[t-N+1, t]$. At certain times $\{t_i\}$, the current parameter set estimate would be used to update the controller.

There are myriad ways to select the reset times $\{t_i\}$. As an illustration, let the reset times be

$$\{t_i\} = \{kN : \begin{array}{l} \text{parameter identifiablity on } [1+(k-1)N, kN], \\ \text{control design exists,} \\ \text{parameter set estimate } \in \Theta_o \} \end{array} \tag{101}$$

Parameter identifiability is essentially an information test on the data and is equvalent to the notion of ***persistent excitation on the observation interval*** $[t_i - N + 1, t_i]$. If there is no identifiablity, then we keep recording the data. Even with good data, there is still the possibility of being unable to use the resulting parameter set estimate because there is no "numerical" solution for the control design, *e.g.,* whenever there is a near pole and zero cancellation in the model set estimate. Finally, the estimated parameter set may contain some parameters which are outside the set of prior information Θ_o. If this occurs, then the "outliers" can be projected back into the prior set, or else continue to record data. A further discussion of projection methods and resetting schemes are dealt with in Goodwin and Sin(1984) and Bai and Sastry(1987).

Solving the least-squares parameter estimation problem on the moving interval results in the estimate at t_i given by

$$\hat{\theta}_i = \left\langle \phi\phi^T \right\rangle_i^{-1} \langle \phi z \rangle_i \tag{102}$$

where ϕ, z are as previously defined, but now $\langle \cdot \rangle_i$ denotes averaging on the moving interval, *i.e.,*

$$\langle x \rangle_i \stackrel{\text{def}}{=} \frac{1}{N} \sum_{t=t_i-N+1}^{t_i} x(t) \tag{103}$$

The parameter identifiability test is to require that the smallest eigenvalue of $\left\langle \phi\phi^T \right\rangle_i$ is larger than some positive number.

Using this scheme, the adaptive control system can be viewed as an iterative mapping in parameter space. Specifically, let $\hat{\Theta}_i$ denote the "new" parameter set estimate at the end of the ith ***reset interval*** $[t_{i-1}, t_i]$, where

$$\hat{\Theta}_i \stackrel{\text{def}}{=} \left\{ \theta \in \Theta_o : \|\theta - \hat{\theta}_i\| \le \hat{\delta}_i \right\} \tag{104}$$

Hence, corresponding to each model set estimate $\hat{\Theta}_i$, there is a parameter estimate $\hat{\theta}_i \in \Theta_o$ and a radius of uncertainty $\hat{\delta}_i$. The parameter set estimator produces the sequence of "new" parameter set estimates

$$\widehat{\Theta}_i, \quad i = 1, 2, \ldots \tag{105}$$

Correspondingly, we have the sequence of "current" parameter set estimates

$$\overline{\Theta}_i = \widehat{\Theta}_{i-1}, \quad i = 1, 2, \ldots \tag{106}$$

where $\widehat{\Theta}_o$ is an initial estimate.

Using the above relations between "new" and "current" parameter set estimates, we may conceptually express the adaptation as the iterative mapping:

$$\widehat{\Theta}_i = \Gamma_i \left\{ \widehat{\Theta}_{i-1} \right\}, \quad i = 1, 2, \ldots \tag{107}$$

The operator $\Gamma_i \{\cdot\}$ is a time-varying operator mapping parameter sets into parameter sets in $\mathbb{R}^p$. In words, the control design rule generates a controller based on the "current" set estimate $\widehat{\Theta}_{i-1}$. This controller is fixed during the ith reset interval $[t_{i-1}, t_i]$. At the end of this interval, the "current" estimate $\widehat{\Theta}_{i-1}$ is replaced by the "new" estimate $\widehat{\Theta}_i$, and the controller is redesigned. Essentially, the "new" estimate is a function of the "current" (and now past) estimate as well as the signals present in the closed-loop system during the ith reset interval. Thus, the operator $\Gamma_i \{\cdot\}$ signifies the relation between the "current" and "new" parameter set estimates. A further discussion and anlysis of the properties of this operator can be found in Phillips, Kosut, and Franklin(1988). It is further revealed there that fixed- points of this operator are in fact equilibriums of the averaged system corresponding to either slow adaptation or least-squares based parameter adaptive algorithms. What we have essentially shown in this paper is that the fixed points can be restricted to a set bounded by *a priori* knowledge about unmodeled dynamics and disturbances.

8 Concluding Remarks

We have presented some theoretical results and simulation experiments which give credence to the use of a parameter set estimator in an adaptive control scheme where the on- line design rule is prepared to accept uncertainties of the same form. The main result shows that with a sufficient but reasonable amount of ***a priori*** information, the parameter set estimator can always capture the true system. To make the range of uncertainty as small as is required by control design specifications will necessitate an appropriate choice of identification experiment design variables, such as data filters, and input spectrum.

Ackowlegements

Inspiration for this work is in part due to a recent visit to Sweden with K.J. Åström and B. Wittenmark of the Lund Institute of Technology, and with L. Ljung at Linköping University.

A Appendix

A.1 Preliminaries

We first state some preliminary definitions and results.

Definition: *The discrete-Fourier-transform, or DFT, of a sequence $\{f\}$ on $t \in [0, N-1]$ is defined as*

$$F_N(\omega) = \mathrm{DFT}\{f\} = \frac{1}{N}\sum_{t=0}^{N-1} f(t)e^{-j\omega t}, \ \ \omega \in \Omega_N$$

where Ω_N are the DFT frequencies

$$\Omega_N = \{\frac{2\pi k}{N} : k \in [0, N-1]\}$$

The DFT admits an inverse transform, or IDFT, such that

$$f(t) = \mathrm{IDFT}\{F_N\} = \sum_{\omega\in\Omega_N} F_N(\omega)e^{j\omega t}, \ \ t \in [0, N-1]$$

Outside of the interval $[0, N-1]$, $f(t) \neq \mathrm{IDFT}\{F_N\}$ in general, unless f is N-periodic. We use the notation f_N to denote the N- periodic *extension* of f which is defined for all $t \in (-\infty, \infty)$ by

$$f_N(t) = \sum_{\omega\in\Omega_N} F_N(\omega)e^{j\omega t}$$

Observe that for $t \in [0, N-1], f_N(t) = f(t)$. An immediate consequnce of the above definitions is

Theorem 4 (Parseval's Theorem)

$$\left\langle xy^T\right\rangle_N = \frac{1}{N}\sum_{t=0}^{N-1} x(t)y^T(t) = \sum_{\omega\in\Omega_N} X_N(\omega)Y_N(\omega)^*$$

The next result shows how DFT's are transformed by stable linear systems. The result as stated is a slight modification of Theorem 2.1 in Ljung(1987), and for this reason a proof is presented. A similar result can also be found in LaMaire *et al.* (1987).

Lemma 2 (Linear Transformation of DFT) *Let $\{y, u : t = 0, \ldots, N-1\}$ denote sequences with DFT's $Y_N(\omega), U_N(\omega)$ for $\omega \in \Omega_N$. Suppose that*

$$y = Gu$$

where G is a stable transfer function whose pulse response sequence $\{g\}$ satisfies

$$|g(t)| \leq M\rho^t, \ \forall t \geq 0$$

for constants $M > 0, \rho \in (0,1)$. Then:

$$|Y_N(\omega) - G(e^{j\omega})U_N(\omega)| \leq \frac{1}{N}\left(\sup_{t\in(-\infty,-1]} |u(t) - u_N(t)|\right)\left(\frac{M\rho(1-\rho^N)}{(1-\rho)^2}\right)$$

where $\{u_N(t)\}$ is the N-periodic extension of $u(t)$, that is,

$$u_N(t) = \sum_{\omega\in\Omega_N} U_N(\omega)e^{j\omega t},\ t \in (-\infty,\infty)$$

If $u(t)$ is N-periodic, then

$$Y_N(\omega) = G(e^{j\omega})U_N(\omega)$$

Proof:

For $t \in [0, N-1]$, we have

$$\begin{aligned}
y(t) &= \sum_{k=-\infty}^{t} g(t-k)u(k) \\
&= \sum_{k=-\infty}^{-1} g(t-k)u(k) + \sum_{k=0}^{t} g(t-k)u_N(k) \\
&= \sum_{k=-\infty}^{-1} g(t-k)u(k) + \sum_{k=-\infty}^{t} g(t-k)u_N(k) - \sum_{k=-\infty}^{-1} g(t-k)u_N(k) \\
&= \sum_{k=-\infty}^{t} g(t-k)u_N(k) + \sum_{k=-\infty}^{-1} g(t-k)[u(k) - u_N(k)]
\end{aligned}$$

From the definition of the DFT,

$$Y_N(\omega) = G(e^{j\omega})U_N(\omega) + R_N(\omega)$$

where

$$R_N(\omega) = \frac{1}{N}\sum_{t=0}^{N-1} e^{-j\omega t}\left(\sum_{k=-\infty}^{-1} g(t-k)[u(k) - u_N(k)]\right)$$

Thus,

$$\begin{aligned}
|R_N(\omega)| &\leq \frac{1}{N}\sup_{t\in(-\infty,-1]} |u(t) - u_N(t)| \sum_{t=0}^{N-1}\sum_{k=-\infty}^{-1} |g(t-k)| \\
&\leq \frac{1}{N}\sup_{t\in(-\infty,-1]} |u(t) - u_N(t)| \sum_{t=0}^{N-1}\sum_{k=-\infty}^{-1} M\rho^{t-k} \\
&= \frac{1}{N}\sup_{t\in(-\infty,-1]} |u(t) - u_N(t)| M\left(\sum_{t=0}^{N-1}\rho^t\right)\left(\sum_{k=-\infty}^{-1}\rho^{-k}\right) \\
&= \frac{1}{N}\sup_{t\in(-\infty,-1]} |u(t) - u_N(t)| M\left(\frac{1-\rho^N}{1-\rho}\right)\left(\frac{\rho}{1-\rho}\right)
\end{aligned}$$

This completes the proof.

From the above Lemma we obtain:

Corllary 2 *Let $\{y, u : t \in [0, N-1]\}$ have respective DFT's $\{Y_N, U_N : \omega \in \Omega_N\}$. Let G be a stable transfer function whose pulse response sequence $\{g\}$ satisfies*

$$|g(t)| \leq M\rho^t, \ \forall t \geq 0$$

for constants $M > 0, \rho \in (0,1)$. Then:

$$\langle y(Gu) \rangle_N = \sum_{\omega \in \Omega_N} Y_N(\omega)[G(e^{j\omega})U_N(\omega)]^* + e_N$$

where

$$|e_N| \leq \frac{1}{N} M \sup_{t \in [0, N-1]} |u(t)| \sup_{t \in (-\infty, -1]} |u(t) - u_N(t)| \left(\frac{1-\rho^N}{1-\rho} \right) \left(\frac{\rho}{1-\rho} \right)$$

with $u_N(t)$ the N-periodic extension of $u(t)$.

A.2 Proof of Theorem 1

We start with the expression for parameter error from Lemma 1, *i.e.*,

$$\tilde{\theta}_N^\Delta = \langle \psi \varepsilon_\Delta \rangle_N = - \langle \psi(W_F \Delta_A y) \rangle_N + \langle \psi(W_F \Delta_B u) \rangle_N$$

where we have defined

$$\psi = \left\langle \phi \phi^T \right\rangle_N^{-1} \phi$$

Applying the previous lemma and corollary to each term gives:

$$\tilde{\theta}_N^\Delta = - \sum_{\omega \in \Omega_N} [W_F(e^{j\omega})\Delta_A(e^{j\omega})Y_N(\omega)]^* \Psi_N(\omega) + \sum_{\omega \in \Omega_N} [W_F(e^{j\omega})\Delta_B(e^{j\omega})U_N(\omega)]^* \Psi_N(\omega) + e_N$$

where

$$\begin{aligned} \Psi_N(\omega) &= \left\langle \phi \phi^T \right\rangle_N \Phi_N(\omega) \\ \|e_N\| &\leq \frac{\beta}{N} \end{aligned}$$

with β as defined in the Theorem. Under the assumptions of the theorem, namely the known frequency bounds on $\Delta_B(e^{j\omega})$ and $\Delta_A(e^{j\omega})$, the indicated bound (37) follows. Ovserve also that

$$\sup\{\| \sum_{\omega \in \Omega_N} [W_F(e^{j\omega})\Delta_A(e^{j\omega})Y_N(\omega)]^* \Psi_N(\omega)\| : |\Delta_A(e^{j\omega})| \leq |W_A(e^{j\omega})|\} = \|\gamma_A\|$$

Thus, the bound (37) is asymptotically (in N) tight. This completes the proof.

References

B.D.O. Anderson, R.R. Bitmead, C.R. Johnson,Jr.,P.V. Kokotovic, R.L. Kosut, I.M.Y. Mareels, L. Praly, and B.D. Riedle (1986), *Stability of Adaptive Systems: Passitivity and Averaging Analysis,* MIT Press, 1986.

K. J. Åströmand B. Wittenmark (1988) *Adaptive Control*, Addison- Wesley, to appear.

E.-W. Bai and S. Sastry (1987), "Global Stability Proofs for Continuous -Time Indirect Adaptive Control Schemes", *IEEE Trans. Aut. Control*, vol. AC-32, pp. 537-543.

J. C. Doyle, J. E. Wall, and G. Stein, (1982), "Performance and Robustness Analysis for Structured Uncertainty", *Proc. IEEE CDC*, orlando, FL, Dec. 1982.

M. Gevers and L. Ljung (1986), "Optimal Experiment Design With Respect to the Intended Model Application", *Automatica*, vol. 22, no. 5, pp. 543 -554.

G. C. Goodwin and M. E. Salgado (1988), "Quantification of Uncertainty in Estimation Using an Embedding Principal", Tech. Report, Univ. of Newcastle , Australia.

G. C. Goodwin and k. S. Sin (1984), *Adaptive Filtering, Prediction , and Control*, Prentice-Hall, New Jersey.

F. Hansen (1988), *A Fractional Representation Approach to Closed- Loop Identification and Experiment Design*, PhD. Dissertation, Dept. of Electrical Engr., Stanford University.

R.L. Kosut (1986), "Adaptive Calibration: An Approach to Uncertainty Modeling and On-Line Robust Control Design," *Proc. 25th IEEE CDC,* Athens, Greece, Dec. 1986.

R.L. Kosut (1987), "Adaptive Uncertainty Modeling: On-Line Robust Control Design," *Proc. 1987 ACC,* Minneapolis, MN, June 1987.

R.L. Kosut (1988), "Adaptive Robust Control via Transfer Function Uncertainty Estimation", *Proc. 1988 ACC*, June 1988, Atlanta, GA.

J.M. Krause and P.P. Khargonekar (1987), "On an Identification Problem Arising in Robust Adaptive Control", *Proc. 26th IEEE CDC*, Los Angeles, CA, Dec. 1987.

R. O. LaMaire, L. Valavani, M. Athans, and G. Stein (1987), "A Frequency- Domain Estimator for Use in Adaptive Control Systems", *Proc. 1987 ACC*, pp 238-244, Minn., MN, June 1987.

L. Ljung (1987), *System Identification: Theory for the User,* Prentice Hall, New Jersey.

I. Mareels, R. Bitmead, M. Gevers, C. R. Johnson, Jr., R. Kosut, and M. A. Poubelle, (1986), "How Exciting Can A Signal Really Be?," *System and Control Letters.*

S. Morrison and B. Walker (1988), "Batch-Least-Squares Adaptive Control in the Presence of Unmodelled Dynamics," *Proc. 1988 ACC*, pp. 774-776, Atlanta, GA, June 1988.

P. J. Parker and R. R. Bitmead (1987), "Adaptive Frequency Response Identification," *Proc. 26th IEEE CDC*, pp. 348-353, Los Angeles, CA, Dec. 1987.

S. M. Phillips, R.L. Kosut, and G.F. Franklin (1988), "An Averaging Analysis of Discrete-Time Indirect Adaptive Control", *Proc. 1988 ACC*, Atlanta, GA, June 1988.

D. Rovner and G. Franklin (1987), "Experiments in Load-Adaptive Control of a Very Flexible One-Link Manipulator," *IEEE Trans. on Aut. Control*, to appear.

M. Vidyasagar, (1984), *Control System Synthesis: A Factorization Approach*, M.I.T. Press, Cambridge, MA.

B. Wahlberg and L. Ljung (1986), "Design variables for Bias Distribution in Transfer Function Estimation", *IEEE Trans. on Aut. Control*, vol AC-31, no. 2, pp. 134-144, Feb. 1986.

B. Wahlberg (1986), "On Model Reduction in System Identification," *Proc. 1986 ACC*, pp. 1260-1266, Seattle, WA, June 1986.

B. Wittenmark (1986), "On the Role of Filters in Adaptive Control", Report EE8662, Dec. 1986, Dept. of Electrical and Computer Engineering, University of Newcastle, Australia.